INTERNATIONAL JOURNAL OF ABSTRACTS

STATISTICAL THEORY AND METHOD

VOLUME 2 · NUMBER 4
1961

PUBLISHED FOR
THE INTERNATIONAL STATISTICAL INSTITUTE
OLIVER AND BOYD

VOLUME 2 · NUMBER 4

COPYRIGHT © 1961 INTERNATIONAL STATISTICAL INSTITUTE

Abbreviation (*World List* Style):
Int. J. Abstr., Statist. Theory

Short Title:
Statistical Theory Abstracts

PRINTED IN GREAT BRITAIN BY
OLIVER AND BOYD LTD., EDINBURGH

INTERNATIONAL JOURNAL OF ABSTRACTS
STATISTICAL THEORY AND METHOD

COVERAGE OF JOURNAL

THE aim of this journal of abstracts is to give complete coverage of papers in the field of statistical theory and new contributions to statistical method. Papers which report only applications or examples of existing statistical theory and method will not be included. There are approximately two hundred and fifty journals published in various parts of the world which are wholly or partly devoted to the field of statistical theory and method and which will be brought within the scope of this journal of abstracts. A complete list of journals covered is printed in the annual Index Supplement. In the case of the following journals, however, being those which are wholly devoted to statistical theory—all contributions, whether a paper, note or miscellanea, will be abstracted :

Annals of Mathematical Statistics
Biometrika
Journal, Royal Statistical Society (Series B)
Bulletin of Mathematical Statistics
Annals, Institute of Statistical Mathematics
Sankhyā (Series A)

Within the larger group of journals, which are not wholly devoted to statistical theory and method, there are some journals which have the vast majority of their contributions in this field. These journals, therefore, will be abstracted on a virtually complete basis :

Biometrics
Metrika
Metron
Review, International Statistical Institute
Technometrics
Sankhyā (Series B)

If any reader of this journal discovers a paper which happens to have been overlooked, the General Editor will be pleased to be informed so that the appropriate abstract can be made : always provided that the date of publication is after 1st October 1958, when the abstracting for this journal commenced.

In addition to the ordinary journals, there are two kinds of publication which fall within the scope of this journal of abstracts. They are the experiment and other research station reports—which occur particularly in the North American region—and the reports of conferences, symposia and seminars. Whilst these latter may be included in the book review sections of journals it is unusual for any individual contribution to be noted at any length. These publications are, in effect, special collections of papers and for this reason the appropriate arrangements will be made for them to be included in this journal. By the same token, abstracts of papers given at conferences and reproduced in an appropriate journal will be disregarded until the definitive publication is available.

FORMAT OF JOURNAL

The abstracts will be up to 400 words long—the recommendation of UNESCO for the " long " abstract service : they will be in the English language although the language of the original paper will be indicated on the abstract together with the name of the abstractor. In addition, the address of the author(s) will be given in sufficient detail to facilitate contact in order to obtain further detail or request an offprint. Suitable indexes will be provided annually in a supplement.

The scheme of classification has been developed upon lines that will facilitate the transfer to punched cards of the code numbers allocated to each abstract: to allow for future development it is suggested that use is made of 4-column fields. Each abstract will have two classification numbers: the primary number in heavy type to indicate the basic topic of the paper and the secondary number in brackets to take account of the most important cross-reference. The sheets of the journal are colour-coded according to the twelve main sections of the classification and it should be noted that it is the main section number of the primary classification which determines the colour code for each abstract. It is believed that this method of colour-coding the pages will provide a distinctive visual aid to the identification of abstracts both when the journal is in bound form or dismantled and filed on cards or in binders. The format and simple binding allows of the following alternative treatments by users of the journal:

(*a*) Leave intact as a shelf-periodical.

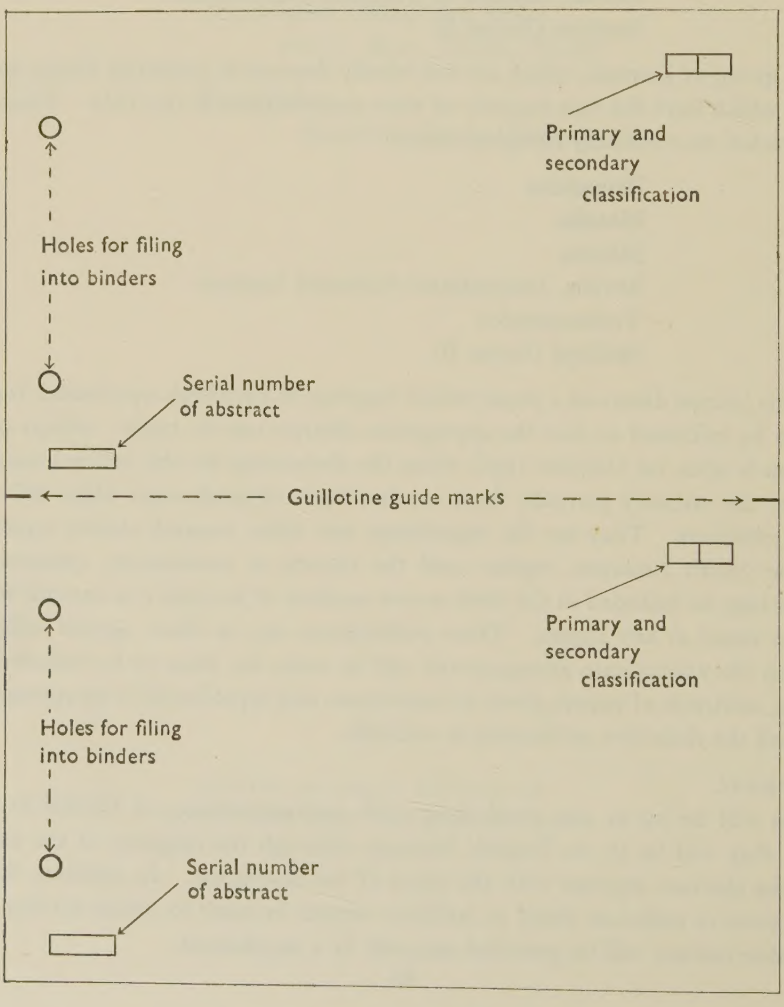
(*b*) Split and filed in page form according to the main sections of the classification.

(*c*) Split and guillotine (single cut) each page ready for:

  (i) pasting onto standard index cards.

  (ii) filing in loose-leaf or other binders—for which the appropriate holes are punched.

It will also be possible for those users who need a completely referenced file to insert skeleton cards or sheets according to the secondary classification number provided on each abstract.


Primary and secondary classification
Holes for filing into binders
Serial number of abstract
Guillotine guide marks
Primary and secondary classification
Holes for filing into binders
Serial number of abstract

JOURNAL OF THE ROYAL STATISTICAL SOCIETY

SERIES A (GENERAL), four issues a year, 25s. per part, annual subscription £4. 4s. (U.S. $12) post free; SERIES B (METHODOLOGICAL), two issues a year, 40s. per part, annual subscription £3. 2s. (U.S. $9) post free.

SERIES A (GENERAL)

VOL. 124, PART 3, 1961

An Experiment in the Marking of an Examination (with Discussion)	D. V. LINDLEY
The Intercorrelation Between the Nutrients Consumed by a Group of Families in the Netherlands	E. F. DRION
Diets of Bank Clerks	J. A. HEADY
Discussion on the Papers by Dr. DRION and MR. HEADY	
Sequential Analysis: a survey	N. L. JOHNSON
A Technique for Estimating a Contingency Table, given the Marginal Totals and Some Supplementary Data	D. FRIEDLANDER
Statistical Victoriana	A. R. ILERSIC
Wholesale Prices in 1960. The Editor of *The Statist*	
Obituary: CHARLES JORDAN	
Reviews; Current Notes; Book List	

SERIES B (METHODOLOGICAL)

VOL. 23, No. 2, 1961

A Test of Homogeneity of Means under Restricted Alternatives (with Discussion)	D. J. BARTHOLOMEW
Optimal Programmes for Varietal Selection (with Discussion)	R. N. CURNOW
Some Stochastic Processes with Absorbing Barriers	T. L. SAATY
Some Sequential Tests using Range	W. G. GILCHRIST
First Emptiness in a Finite Dam	B. WEESAKUL
The Construction of Optimal Designs for the One-Way Classification Analysis of Variance	J. AITCHISON
Some Simple Wear-Dependent Renewal Processes	A. MERCER
An Exact Distribution of the Fisher-Behrens-Welch Statistic for Testing the Difference between the Means of Two Normal Populations with Unknown Variances	W. D. RAY and A. E. N. T. PITMAN
Closed Queueing Systems: A Generalisation of the Machine Interference Model	F. BENSON and G. GREGORY
Testing for a Jump in the Spectral Function	E. J. HANNAN
Moments of a Truncated Bivariate Normal Distribution	S. ROSENBAUM
On Quadratic Estimates of the Interclass Variance for Unbalanced Designs	R. R. READ
Mixtures of Geometric Distributions	H. E. DANIELS
Prediction by Exponentially Weighted Moving Averages and Related Methods	D. R. COX
The Randomisation Bases of the Problem of the Amalgamation of Weighted Means	D. E. BARTON and C. L. MALLOWS
Estimation from Quantiles in Destructive Testing	H. P. STOUT and F. STERN
A Note on Maximum-Likelihood in the Case of Dependent Random Variables	S. D. SILVEY
Some Properties of the Asymptotic Variance of the Sample Quantiles and Mid-Ranges	P. K. SEN
Similarities between Likelihoods and Associated Distributions a Posteriori	D. A. SPROTT
A Method of Maximum Likelihood Estimation	F. S. G. RICHARDS
The Intervals between Regular Events Displaced in Time by Independent Random Deviations of Large Dispersion	T. LEWIS
An Approach to the Scheduling of Jobs on Machines	E. PAGE

ROYAL STATISTICAL SOCIETY, 21 BENTINCK STREET, LONDON, W.1

THE ANNALS OF MATHEMATICAL STATISTICS

Vol. 32, No. 4—December 1961

TABLE OF CONTENTS

An Approach to Time Series Analysis	Emanuel Parzen
Some Model I Problems of Selection	E. L. Lehmann
Bayes Rules for a Common Multiple Comparisons Problem and Related Student-t Problems	David B. Duncan
The Use of Least Favorable Distributions in Testing Composite Hypotheses	H. E. Reinhardt
Asymptotic Efficiency in Polynomial Estimation	Paul G. Hoel
Maximum Likelihood Estimation of a Linear Functional Relationship	C. Villegas
Sequential x^2—and T^2—Tests	Edward J. Jackson and Ralph A. Bradley
Estimating the Parameters of Negative Exponential Populations from One of Two Order Statistics	H. Leon Harter
On the Two Sample Problem: A Heuristic Method for Constructing Tests	V. P. Godambe
A Nonparametric Test for the Problem of Several Samples	V. P. Bhapkar
Distribution of the Anderson-Darling Statistic	Peter A. W. Lewis
Errors in Discrimination	S. John
On the Monotonic Character of the Power Functions of Two Multivariate Tests	S. N. Roy and W. F. Mikhail
The Moments of Elementary Symmetric Functions of the Roots of a Matrix in Multivariate Analysis	Tito A. Mijares
Variance Components in the Unbalanced 2-Way Nested Classification	S. R. Searle
Some Main-Effect Plans and Orthogonal Arrays of Strength Two	Sidney Addelman and Oscar Kempthorne
On a Geometrical Method of Construction of Partially Balanced Designs with Two Associate Classes	Esther Seiden
On Some Methods of Construction of Partially Balanced Arrays	I. M. Chakravarti
Some Further Designs of Type O:PP	G. H. Freeman
Sufficiency in the Undominated Case	D. L. Burkholder
On a Special Class of Recurrent Events	M. P. Schützenberger
Maximum Likelihood Characterisation of Distributions	Henry Teicher
On the Distribution of First Significant Digits	Roger S. Pinkham
Markov Renewal Processes: Definitions and Preliminary Properties	Ronald Pyke
Markov Renewal Processes with Finitely Many States	Ronald Pyke
A Convexity Property in the Theory of Random Variables Defined on a Finite Markov Chain	H. D. Miller
Limit Distribution in the Theory of Counters	G. Sankaranarayanan
The Transient Behavior of a Single Server Queuing Process with Recurrent Input and Gamma Service Time	Lajos Takács
Efficient Estimation of a Regression Parameter for Certain Second Order Processes	Charlotte T. Striebel
Two Similar Queues in Parallel	J. F. C. Kingman
Queues with Batch Departures I	F. G. Foster and K. M. Nyunt
Notes	
On the Chapman-Kolmogorov Equation	Jack Karush
A Generalisation of a Theorem of Balakrishnan	N. Donald Ylvisaker
The Opinion Pool	M. Stone
Correction Notes	
Abstracts of Papers	
News and Notices	
Report of the President for 1961	
Report of the Editor for 1961	
Publications Received	

BIOMETRIKA

Volume 48, Parts 3 and 4 **CONTENTS** December 1961

Memoirs

DARROCH, J. N. The two sample capture-recapture census when tagging and sampling are stratified.

FRASER, D. A. S. The fiducial method and invariance.

WETHERILL, G. B. Bayesian sequential analysis.

BROWN, BYRON WM. (Jr.). Some properties of the Spearman estimator in bioassay.

SLATER, PATRICK. Inconsistencies in a schedule of paired comparisons.

MORGENTHALER, GEORGE W. Some circular coverage problems.

BARTHOLOMEW, D. J. Ordered tests in the analysis of variance.

YOUNG, D. H. Quota fulfilment using unrestricted random sampling.

WALKER, A. M. Large-sample estimation of parameters for moving-average models.

REIERSØL, OLAV. Linear and non-linear multiple comparisons in logit analysis.

SAW, J. G. Estimation of the normal population parameters given a type I censored sample.

QUESENBERRY, C. P. and DAVID, H. A. Some tests for outliers.

KINGMAN, J. F. C. The ergodic behaviour of random walks.

KSHIRSAGAR, A. M. The goodness-of-fit of a single (non-isotropic) hypothetical principal component.

JOHN, S. On the evaluation of the probability integral of the multivatiate t-distribution.

IMHOF, J. P. Computing the distribution of quadratic forms in normal variables.

CHANDA, K. C. Comparative efficiencies of methods of estimating parameters in linear autoregressive schemes.

MOSTELLER, FREDERICK and YOUTZ, CLEO. Tables of the Freeman-Tukey transformations for the binomial and Poisson distributions.

Miscellanea: Contributions by J. D. BIGGERS; E. J. BURR and GWENDA CANE; C. F. CROUSE; P. D. FINCH; GERALD J. GLASSER and ROBERT F. WINTER; D. HOGBEN, R. S. PINKHAM and M. B. WILK; B. K. KALE; J. G. SAW; A. ZINGER.

Reviews.

Other Books received.

Corrigenda.

The subscription, *payable in advance*, is 54/- (or $8.00) per volume (including postage). Cheques should be made payable to *Biometrika*, crossed " a/c *Biometrika Trust* " and sent to The Secretary, Biometrika Office, University College, London, W.C.1. All foreign cheques must be drawn on a Bank having a London agency.

Issued by THE BIOMETRIKA OFFICE, University College, London

BIOMETRICS

Journal of the Biometric Society

VOL. 17, NO. 4 DECEMBER 1961

CONTENTS

The Poisson Pascal Distribution	*S. K. Katti and John Gurland*
Some Rank Sum Multiple Comparison Tests	*Robert G. D. Steel*
The Estimation of Repeatability and Heritability from Records Subject to Culling	*R. N. Curnow*
Three Classes of Univariate Discrete Distributions	*C. G. Khatri and I. R. Patel*
Further Consideration of Methodology in Studies of Pain Relief	*Paul Meier and Spencer M. Free, Jr.*
Fitting a Geometric Progression to Frequencies	*E. J. Williams*
Computing Procedures for Estimating Components of Variance in the Two-way Classification	*S. R. Searle and C. R. Henderson*
Optimum Sample Size in Animal Disease Control	*A. W. Nordskog and H. T. David, H. B. Eisenberg*
Numerical Aspects of the Regression of Offspring on Parent	*H. E. McKean and B. B. Bohren*
Some Hypotheses Concerning Two Phase Regression Lines	*P. Sprent*
QUERIES AND NOTES	
A Practical Application of a Theoretically Inefficient Design	*C. P. Cox*
Query: On a Graphical Sequential Test	*I. D. J. Bross*
A Simple Method of Fitting the Regression Curve $y=\alpha+\delta x+\beta\rho^x$	*B. K. Shah*
Query: On a Follow-up Study of the Growth of Children	*M. R. Sampford*
BOOK REVIEW	
H. De Jonge: Quantitative Methods in Pharmacology	*D. J. Finney*

Biometrics is published quarterly. Its objects are to describe and exemplify the use of mathematical and statistical methods in biological and related sciences, in a form assimilable by experimenters. The annual non-member subscription rate is $7. Inquiries, orders for back issues and non-member subscriptions should be addressed to:

BIOMETRICS, Department of Statistics, The Florida State University, Tallahassee, Florida

SANKHYĀ

The Indian Journal of Statistics

Contents of Vol. 23, Series A, Part 3

An elementary theory of the Brownian motion random function	J. Kampé De Fériet
On testing the mean of a discrete linear process	K. R. Parthasarathy
On a problem of Bartholomew in life testing	P. S. Swamy and S. A. D. C. Doss
Studentisation of two-stage sample means from normal populations with unknown common variance	Harold Ruben
Testing for the single outlier in a regression model	K. S. Srikantan
Minimal sufficient statistics for the balanced incomplete block design under an Eisenhart model II	David L. Weeks and Franklin A. Graybill
Asymptotic bias and variance of ratio estimates in generalised power series distributions and certain applications	G. P. Patil
Method of constructing two mutually orthogonal latin squares of order $3n+1$	P. Kesava Menon
Combinatorial arrangements analogous to orghogonal arrays	G. Radhakrishna Rao
On the structure and combinatorial properties of certain semi-regular group divisible designs	W. S. Saraf
A factorial approach to construction of true cost of living index and its application in studies of changes in national income	K. S. Banerjee
Generation of random permutations of given number of elements using random sampling numbers	C. Radhakrishna Rao

The subscription price per volume for each of the series is: U.S.A. and Canada: $10.00; Other Countries: £3. The subscription price per volume for both the series taken together is: U.S.A. and Canada: $18.00; Other Countries: £5, 6s.

Subscription is payable in advance to: The Secretary, Statistical Publishing Society, 204/1 Barrackpore Trunk Road, Calcutta-35, India.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Volume 56 December 1961 Number 296

RECTIFYING INSPECTION OF LOTS	F. J. Anscombe
ON COMPARING INTENSITIES OF ASSOCIATION BETWEEN TWO BINARY CHARACTERISTICS IN TWO DIFFERENT POPULATIONS	Agnes Berger
OCCUPATIONAL COMPONENTS OF EDUCATIONAL DIFFERENCES IN INCOME	Otis Dudley Duncan
FAILURE OF ENUMERATORS TO MAKE ENTRIES OF ZERO: ERRORS IN RECORDING CHILDLESS CASES IN POPULATION CENSUSES	M. A. El-Badry
NOTE ON THE MISSING PLOT PROCEDURE IN A RANDOMIZED BLOCK DESIGN	Leroy Folks and Del Lon West
A BIVARIATE EXTENSION OF THE EXPONENTIAL DISTRIBUTION	John E. Freund
STEPWISE LEAST SQUARES: RESIDUAL ANALYSIS AND SPECIFICATION ERROR	Arthur S. Goldberger
STATISTICAL METHODS FOR THE MOVER-STAYER MODEL	Leo. A. Goodman
GAMMA DISTRIBUTION IN ACCEPTANCE SAMPLING BASED ON LIFE TESTS	Shanti S. Gupta and Phyllis A. Groll
ON THE RESOLUTION OF STATISTICAL HYPOTHESES	Robert V. Hogg
A SIMPLE THEORETICAL APPROACH TO CUMULATIVE SUM CONTROL CHARTS	Norman L. Johnson
A NOMOGRAPH FOR COMPUTING PARTIAL CORRELATION COEFFICIENTS	Ruth W. Lees and Frederic M. Lord
NON-ADDITIVITY IN TWO-WAY ANALYSIS OF VARIANCE	John Mandel
THE STATISTICAL ANALYSIS OF INDUSTRY STRUCTURE: AN APPLICATION TO FOOD INDUSTRIES	Lee E. Preston and Earl J. Bell
THE ASYMPTOTIC VARIANCES OF METHOD OF MOMENTS ESTIMATES OF THE PARAMETERS OF THE TRUNCATED BINOMIAL AND NEGATIVE BINOMIAL DISTRIBUTIONS	S. M. Shah
FORECASTING INDUSTRIAL PRODUCTION	H. O. Stekler
RESIDENCE HISTORIES AND EXPOSURE RESIDENCES FOR THE UNITED STATES POPULATION	Karl E. Taeuber, William M. Haenszel and Monroe G. Sirken
TESTING THE INDEPENDENCE OF REGRESSION DISTURBANCES	Henri Theil and A. L. Nagar

FOR FURTHER INFORMATION, PLEASE CONTACT

AMERICAN STATISTICAL ASSOCIATION, 1757 K STREET, N. W., WASHINGTON 6, D. C.

AUTHOR INDEX

VOL. 2: No. 4 (599-780)

Abt, K. 7.6

Balogh, T. 10.8
Banerjee, S. K. 4.4
Bánkövi, G. 11.7
Barton, D. E. 1.1
Barlow, R. E. 10.1
Basmann, R. L. 6.6, 6.6
Batschelet, E. 6.9
Baxter, G. 1.0
Bazley, N. W. 11.7
Beatty, G. H. 4.5
Berkson, J. 4.1, 4.1
Bhapkar, V. T. 7.1
Bhat, B. R. 10.2
Birnbaum, A. 4.3
Birnbaum, Z. W. 1.1
Blom, G. 10.1, 10.1
Bos, G. G. J. 10.4
Box, G. E. P. 9.3
Bradley, R. A. 9.1
Breiman, L. 1.6
Brooks, S. H. 9.3
Brown, B. W. 6.8
Burrows, G. L. 8.8

Candler, W. 0.8
Chorneyko, I. 4.7
Chow, G. C. 6.1
Christ, C. F. 6.6
Ciucu, G. 10.1
Cohen, A. C., Jr. 4.3, 4.3
Cohn, R. 3.8
Cornish, E. A. 4.8
Corsten, L. C. A. 9.1
Cox, D. R. 8.8, 10.1
Csáki, P. 6.2, 6.2

Daniel, C. 9.2
Danskin, J. N. 0.7, 10.5
David, Florence N. 1.1
Davies, O. L. 8.8
Davies, P. J. 11.7
Deeming, Lola S. 11.9, 11.9, 11.9, 11.9
Dinkelbach, W. 0.8
Doig, Alison G. 0.8
Draper, N. R. 9.3, 9.3

Ehrenfeucht, A. 1.5
Elteren, P. van 5.7
Elveback, Lillian 4.1
Epstein, B. 2.6, 4.4, 4.4, 5.2
Erdös, P. 1.3
Esary, J. D. 1.1
Evans, Marilyn J. 11.1

Fend, A. V. 4.1
Ferschl, F. 0.8
Finch, P. D. 0.6, 10.4
Fischer, J. 6.2, 6.2
Fisz, M. 1.5, 10.0
Fix, Evelyn 1.1
Fortet, R. 0.0
Freimer, M. 2.1

Ghellinck, G. de 1.8
Ghosal, A. 10.4
Godambe, V. P. 4.9
Gold, B. 2.1
Goldsmith, P. L. 8.9
Golouin, N. E. 2.5
Goodman, L. A. 4.3
Gower, J. C. 10.9
Gladysz, S. 1.5
Graybill, F. A. 9.1
Gupta, S. S. 8.8
Gyires, B. 1.5

Hald, A. 8.8
Harter, H. L. 4.4
Haight, F. A. 10.4, 11.9
Halperin, M. 8.8
Healy, W. C., Jr. 4.3
Hellmich, K. 0.8
Hildreth, C. 6.6
Hill, G. W. 11.1
Hoffman, A. J. 9.1
Hogg, R. V. 3.6
Houthaker, H. S. 0.8
Hunter, J. S. 9.3
Hunter, L. C. 10.1

Iosifescu, M. 10.1

Jambunathan, M. V. 8.3
John, S. 2.9
Johnson, E. A. 6.8

Karlin, S. 1.0
Kishen, K. 9.2
Klein, L. R. 6.6
Kloek, T. 6.6
Klotz, J. 5.6
Kosambi, D. D. 0.4
Koshal, R. S. 8.3
Kotlarski, I. 2.9
Kraft, C. A. 1.1
Kramer, C. Y. 9.1
Künzi, H. P. 0.0

Lamperti, J. 10.1
Land, Ailsa H. 0.8
Lehman, E. J., Jr. 2.4
Leslie, P. H. 10.9
Liu, T. C. 6.6
Lukaszewicz, J. 10.4

Mann, H. B. 6.1
Martin, H. 8.8
Meerendonk, H. W. van der 10.4
Mendenhall, W. 2.4
Mennes, L. B. M. 6.6
Mickey, M. R. 9.3
Mikhail, W. F. 9.1
Mosteller, F. 3.8
Murthy, M. N. 4.1, 4.9

Nagar, A. D. 11.7
Nanjamma, N. S. 4.1, 4.9
Narayana, T. V. 4.7
Nelson, E. 10.1
Newman, A. E. 8.1

Olds, E. G. 8.9
Onicescu, O. 1.1

Page, E. S. 8.9
Palásti, Ilona 1.4
Parthasarathy, K. R. 10.6
Patil, G. P. 2.5
Pavate, M. V. 9.1
Pfanzagl, J. 5.0
Plackett, R. L. 4.3
Pleszczyńska, Elzbieta 8.8
Pratt, J. W. 1.1, 3.8
Pratt, J. W. 1.1, 3.8
Proschan, F. 1.0

Raghavarao, D. 9.1, 9.8
Rao, M. M. 10.2
Rao, P. V. 9.1
Raja Rao, B. 8.6
Rényi, A. 1.3
Rider, P. R. 3.8
Robertson, W. H. 6.9
Roy, J. 5.8

Sarkadi, K. 3.8
Sandee, J. 6.0
Sathe, Y. S. 4.7
Saunders, S. C. 1.1
Schneeberger, H. 6.1
Schwarz, G. 9.2
Schweitzer, A. 4.3
Sen, P. K. 8.4
Seidienberg, A. 1.1
Seshadri, V. 9.1
Sethi, V. K. 4.9

Severo, N. C. **8.1**
Shah, B. V. **9.1, 9.2**
Shapiro, J. M. **1.0, 1.5**
Siddiqui, M. M. **3.8, 6.1**
Steffens, F. **0.8**
Stein, C. **4.1**
Störmer, H. **4.4**
Strauch, H. **8.9**
Sukhatme, B. V. **8.3**

Tatsuoka, M. **3.8**
Theil, H. **4.3, 11.8, 11.8**
Theodorescu, R. **10.1, 10.1, 10.1**
Thomasian, A. J. **10.5**
Tikkiwal, B. D. **8.3**
Tintner, G. **0.8**
Trotter, H. F. **1.5**
Trybula, S. **4.0**

Ullrich, M. **1.5**
Urbanik, K. **1.5**
Uzawa, H. **5.6**

Varberg, D. **10.1**
Vartak, M. N. **9.1**
Venekamp, P. E. **11.0**
Votaw, D. F., Jr. **0.8**
Vries, M. de **4.0**

Walker, A. M. **10.7**
Walpole, R. E. **9.1**
Walsh, J. E. **4.6**
Weissberg, A. **4.5**
Wetherill, G. B. **8.8**
Whitfield, H. **8.9**
Wijsman, R. A. **5.7**
Working, H. **10.7**
Wormleighton, R. **4.7**
Wolff, P. M. de **1.2**

Zajta, A. **5.3**
Zubrzycka, Ludmilla **8.0**
Zubrzycki, S. **10.2**

Volume 2. CORRECTIONS

2/447 **1.4** Palásti	Author's christian name is, Ilona.
Author Index: Vol. 2, No. 1	Filippo, E; delete and add, Emmanuelli, F.
,, ,, ,, No. 2	Palásti; delete initial and insert the name, Ilona.

Note The symposium from which papers are reported at **2**/500, 501-2, 503, 510 and 545 was held at Leyden and published in Amsterdam (North Holland Publishing Co.)

The Congress from which papers are reported at **2**/433, 442, 444, 445, 447, 451, 452, 453, 454, 471, 479, 482, 491, 492, 511, 517, 518, 527, 558, 571, 576, 579-80, 587, 592, 741 and 767 was held at Budapest and published there by Akadémiai Kiadó.

CANDLER, W. (University of New England, Australia) **0.8** (0.7)

A " short-cut " method for the complete solution of game theory and feed-mix problems—*In English*

Econometrica (1960) **28**, 618-634 (6 references, 3 tables, 4 figures)

The author states that this " short-cut " method is an extension of the graphical short-cut methods presented by Waugh & Burrows [*Econometrica* (1955) **23**, 18-29] and by Boles [*Journal of Farm Economics* (1956) **38**, 981].

Commencing by defining the terms in the title of his paper, the author remarks that he is using the term " complete solution " to imply that with this method it is not necessary to employ the simplex method after the completion of the " short-cut " method. Even in the problem where one is dealing with a large number of columns and rows, a solution can be reached by use of the " short-cut ". Two-person zero-sum games are used in the discussion of game theory problems: programming problems with all positive resource supplies and with either positive or negative prices are used in the feed-mix problems.

Following a description of the notation employed, the short-cut method is discussed and illustrated by an example dealing with game theory. Concise tables are given both for computing the successive steps in the " short-cut " solution to the problem posed and to illustrate the algebraic representation of this table. After a discussion of strategy, an outline is given of various difficulties which may be encountered and the manner in which strategies can be both improved and eliminated. The feed-mix problem used to illustrate the use of this method is then described and its solution discussed.

In conclusion the author remarks that it is now apparent that, at its present level of development, the short-cut programme appears to require a higher level of decision-making than is usual in the case where the simplex method is employed. He considers therefore, that the simplex method is still superior as a method for use by clerical assistants and high speed computors. The short-cut method can usefully be employed when computers are not available.

The author's main object in writing this paper has, he states, been to develop a complete non-simplex method for the solutions of a wide range of problems in the field of programming: further research is intended into short-cut methods both of general and parametric programming problems.

2/599

(W. R. Buckland)

DINKELBACH, W. & STEFFENS, F. (Köln University) **0.8** (–.–)

The solution of some decision problems by mixed integer linear programming—*In German*

Unternehmensforschung (1961) **5**, 3-14 (7 references)

In addition to usual quantitative restrictions in linear programming the following qualitative restrictions are considered. Certain subsets only of a set of n variables can be non-zero in the final solution of the linear programme.

The following four models are formulated as mixed integer linear programming problems:

(i) The final solution consists of at most exactly one variable.

(ii) The final solution consists of at most exactly k ($k<n$) variables.

(iii) Each activity P_i is divided in n_i sub-activities P_{ir}; the corresponding levels of the activities are x_i and x_{ir}; $i = 1, \ldots, n$, $r = 1, \ldots, n_i$. The final solution contains at most one x_{ir} of P_i; the total number of variables x_{ir} is at most k ($k<n$).

(iv) The set of variables is divided in two subsets. The final solution contains only variables of one of these subsets.

It is proposed to solve the models by the method of Gomory, see " An algorithm of the mixed integer problem " [*RAND Corp. Papers* (1960) p. 1885]. This method is explained in the second part of the paper. An example of production planning demonstrates how the solution of a linear programming is modified by qualitative restrictions as given in the third model.

(W. Piesch)

2/600

DANSKIN, J. M. (Rutgers University, New Jersey, USA) **0.7** (10.5)

Calculus of variations and operations research—*In English*

ARL Technical Note 60-151 (1960), Aero. Res. Lab., Wright-Patterson Air Force Base.
iv+19 pp. (7 references)

The author discusses the role of mathematics, especially the calculus of variations, in operational research, and exhibits some problems of operational research. The first problem is that of distributing a given amount of available resources among various alternatives so as to obtain the maximum return. If the number n of alternatives is finite, the problem may be posed as follows: maximise $F(x_1, \ldots, x_n)$ subject to $\Sigma x_i = X$, $x_i \geqq 0$; where x_i/X denotes the proportion of the resources assigned to the ith alternative. If F is strictly convex, the solution is trivial; one selects the largest of the corner values $F(1, 0, \ldots, 0)$, $F(0, 1, 0, \ldots, 0)$, …, $F(0, 0, \ldots, 1)$. If $F(x_1, \ldots, x_n) = \Sigma f_i(x_i)$, where the f_i are strictly concave, a unique solution is given. If the f_i are concave, but not strictly concave, the solution is no longer unique. A solution is given for the case in which the f_i are S-shaped; first convex, then concave. These solutions of the finite problem are generalised by the use of calculus of variations to solutions of the analogous infinite problem; stated as follows: maximise $\int f[x(t), t]\, dt$ subject to $\int x(t)\, dt = X$, $x(t) \geqq 0$, where the x's are measurable functions of t, with $f(x, t)$ and $f_x(x, t)$ continuous and $f_x(x, t) > 0$.

Examples of the application of the results are given, together with game theory and topology, to the solution of various problems. Many military problems amount to zero-sum, two-person games in which each side disposes his forces according to some distribution and then awaits events. The result is embodied in a payoff function $\pi(x, y)$. The game is said to have a solution if there exist a value v_0 and optimal strategies x^0 and y^0 such that $\pi(x, y^0) \geqq v_0$ for all x and $\pi(x^0, y) \geqq v_0$ for all y. Military problems solved by these methods include: firstly, a stockpiling problem in which it is desired to stockpile a certain item for a contingency, for example war, the latter happening only once;

continued

DANSKIN, J. M. (Rutgers University, New Jersey, USA) **0.7** (10.5) *continued*

Calculus of variations and operations research—*In English*

ARL Technical Note 60-151 (1960), Aero. Res. Lab., Wright-Patterson Air Force Base.
iv+19 pp. (7 references)

secondly, a two-machine-gun duel problem in which a fighter closes on a bomber at a constant rate, so that the rate of fire of his machine gun may be described as a function $x(r)$ of the range. Similarly the bomber has a machine gun, with rate $y(r)$. The guns have limited total amounts of ammunition, hence $\int x(r)\, dr = X$, $\int y(r)\, dr = Y$; thirdly, a problem in which each side places military forces linearly, whose value at point t is $v(t)$. At each point the greater force wins. Similar games have mixed strategies; since if the weaker player discloses his strategy he loses all. Suggestions are given for solving nontrivial min-max problems in which player x chooses his strategy and player y, knowing x's strategy, chooses his own to minimise the payoff to x.

(H. L. Harter)

FERSCHL, F. (Bundeskammer der Gewerblichen Wirtschaft, Vienna) **0.8** (-.-)

A calculation scheme for the simplex method—*In German*

Unternehmensforschung (1961) **5**, 32-41 (2 tables, 1 diagram)

The author starts by describing the normal form of linear programming, generally known as the ordinary maximisation problem. Thereafter a detailed calculation scheme is given to solve the problems mentioned. It differs from the usual ones, as given for instance by Charnes, Cooper and Henderson [*An Introduction to Linear Programming* (1953) New York: Wiley] by an auxiliary calculation scheme which facilitates the computation and the storage of provisional results. The calculation is carefully explained by a flow diagram. It does not make use of matrix calculus and may be applied by non-mathematicians. Two examples illustrate the scheme.

In the appendix the author mentions some methods by which the number of computing errors can be reduced and by which some more general linear programmes can be transformed into those of the normal form.

(W. Dinkelbach)

2/603

FINCH, P. D. (London School of Economics) **0.6** (10.6)

On the covariance determinants of moving average and autoregression models—*In English*

Biometrika (1960) **47**, 194-195 (2 references)

The weakly stationary process defined by

$$\sum_{r=0}^{p} a_r x_{t-r} = \sum_{r=0}^{q} b_r \varepsilon_{t-r}$$

where the $\{\varepsilon_t\}$ are uncorrelated errors of unit variance, has the covariance determinant $X_n = \| \mathscr{E}(x_s x_t) \|$ for $x_0, \ldots, x_n$. The author finds $\lim X_n$ as $n \to \infty$ in terms of the roots of $\Sigma\, a_r z^r = 0$ and $\Sigma\, b_r z^r = 0$ and shows it to have the same value when the a's and b's are interchanged.

The result is obtained by expressing $\mathscr{E}(x_s x_t)$ in terms of the spectral density function of the process, noting that these may be regarded as Toeplitz forms and hence applying a theorem of Grenander & Rosenblatt on the limit of the determinant of such forms.

(D. E. Barton)

2/604

FORTET, R. (Sorbonne University, France) **0.0** (0.8)
Boolean algebra and its applications in operational research—*In French*
Cahiers Cent. Études Recherche Opérat. (1959) **1**, (10 references)

The author introduces the Boolean algebra U consisting of the elements 0 and 1 and the operations of disjunction, conjunction and negation. A variable taking its values in U is called a Boolean variable. A function f of n independent Boolean variables $x_1, x_2, \ldots, x_n$, $f(U)$ is called a Boolean function. Generally a Boolean problem is to find amongst the 2^n sets of values of $(x_1, x_2, \ldots, x_n)$ those that satisfy a given condition C. Two special cases of the condition C are treated:

(i) C is an algebraic equation: $P(x_1, x_2, \ldots, x_n) = 0$, P being an entire algebraic function, that is a function of the form

$$P = \sum_i \alpha_i \prod_{i_i k_i} x_{i_i} x'_{k_i}$$

where Σ is taken in the ordinary, non-Boolean sense, Π denotes Boolean multiplication, the α_i are constants and the elements of U are interpreted as ordinary numbers.

(ii) C is a Boolean equation: $f(x_1, x_2, \ldots, x_n) = 0$, f being a Boolean function.

Methods of solution are discussed with the aid of simple examples and some practical applications are given; firstly to the four-colour problem, secondly to a machine problem and also to the problem of finding elementary Hamiltonian circuits in a given graph.

Finally, the problem of finding a set of values which satisfies C and which minimises a given entire algebraic function is discussed and the author illustrates this with an α problem from the field of operational research. See also abstract No. **2**/430, **0.0**.

(F. W. Steutel)

2/605

HELLMICH, K. (Graz) **0.8** (-.-)
The route of shortest length or duration—*In German*
Math. Tech. Wirtschaft (1960) **4**, 166-175 (7 references, 1 table)

The author gives a solution of the " travelling salesman problem ": n given cities are considered to be connected by a route of shortest length ending at its starting point. The method of solving this problem given by the author can also be used for asymmetrical matrices of the distances. Moreover, the parts of the route given in advance can be taken into account.

The method is illustrated by a numerical example. Firstly, a known method of reduction of the matrix is demonstrated, leading to groups of neighbouring cities, which are each connected by a closed route. Thereby the reduced matrix is divided in submatrices, called cycle matrices and distance matrices: the cycle matrices represent the preliminary subcycles obtained by the reduction. The distance matrices contain all possible distances between cities of different subcycles.

A systematic procedure is given to connect the subcycles to one shortest route by appropriate solution of distances. For this purpose all distances are ordered. Beginning with the shortest distance subsets are systematically formed and investigated; whether they can be used for a shortest cycle or not. The given numerical example is computed by the selection procedure and a computing programme is given for the general case.

(F. Ferschl)

2/606

HOUTHAKER, H. S. (Stanford Univ., California, USA) **0.8** (-.-)
The capacity method of quadratic programming—*In English*
Econometrica (1960) **28**, 62-96 (6 references, 3 figures)

" Quadratic programming, as considered in this paper, is the art of finding non-negative values of several variables $x_1, \ldots, x_n$ which maximise a polynomial of the second degree $\phi(x_1, \ldots, x_n)$ while satisfying given constraints. The latter constraints will be assumed to be linear." After discussing the facts that quadratic programming is of interest in economics because of its flexibility and realism and does not assume, as does linear programming, that the unit cost of each activity or process and the availabilities of those inputs or requirements for those outputs subject to constraint, are all fixed, the author states that possibilities exist by postulating appropriate linear dependencies, to expand linear programming into quadratic programming. In this paper a method is presented for assigning non-negative values to several variables so as to maximise a given quadratic objective function which satisfies given linear constraints with non-negative coefficients. The method used by the author is parametric and iterative; it leads to a solution in a finite number of steps.

Computational aspects are dealt with at some length and include the opinion that, " another problem on which more experience is needed is that of numerical accuracy, particularly the effect of round-off errors ".

Finally the author says it will be interesting to compare the efficiency of the capacity method with that of other methods of quadratic programming.

An appendix gives computing directions for quadratic programming in the case of non-negative constraints.

(W. R. Buckland)

2/607

KOSAMBI, D. D. (Poona University) **0.4** (0.1)
An application of stochastic convergence—*In English*
J. Indian Soc. Agric. Statist. (1959) **11**, 58-72 (7 references)

The author considers the sequence

$$a_x = \sum_{n \leq x} n^{-\sigma} - \sum_{p \leq x} p^{-\sigma} \log p$$

where the first summation is over all integers and the second is over the primes and where σ is a fixed constant. Probabilistic techniques are used to prove that $\underset{x\to\infty}{\text{limit}}\ a_x$ exists for $\sigma > \frac{1}{2}$. From this the author deduces that if a Riemann zeta function is defined by the equation

$$\zeta(s) = \sum_{1}^{\infty} n^{-s} \quad (s = \sigma + it;\ \sigma > 1)$$

and that if it is continued analytically, it has all its non-trivial zeros on the vertical line $\sigma = \frac{1}{2}$.

(S. R. S. Varadhan)

2/608

KÜNZI, H. P. (Universität Zürich) **0.8** (-.-)

Non-linear programming—*In German*

Ablaufs Planungsforschung (1961) **2**, 8-14

This article is a paper presented at a recent meeting in Münster of the AKOR. The author starts with a short review on linear programming: he explains the mathematical problems and the ideas of two methods of solution, the simplex-method of Dantzig and the gradient-method of Frisch.

A problem in production planning in which the prices depend on the output is given as an example of a non-linear programme. The problem of the quadratic programme is formulated. Before discussing the methods of solution for a quadratic programme the theorem of Kuhn-Tucker is given, which concerns the extrema of convex functions under convex side conditions for non-negative variables. The necessary and sufficient conditions of the Kuhn-Tucker theorem can be applied in the special case of the quadratic programming provided the quadratic objective function is semi-definite. The methods of solution of Barankin and Dorfman are based on this theorem. Barankin's method is explained in detail and some connections with the simplex-method are pointed out. Afterwards Beale's method [*Proc. Camb. Phil. Soc.* (1954) **50**, 513-523] is considered; this also follows the simplex-method. Beale starts with a basic solution, but uses variables without restrictions on their signs.

Finally the paper gives a short survey on other methods of non-linear programming. Houthakker's capacity-method [*Econometrica* (1960) **28**, 62-96; abstracted in this journal No. **2**/607, **0.8**] is fully discussed and, furthermore, is a method which starts with an unrestricted extremum and varies this by adding side-conditions successively: the gradient methods are quoted.

Note: The references to this paper are to appear in a later issue of *Ablaufs Planungsforschung*.

(F. Ferschl)

2/609

LAND, Ailsa H. & DOIG, Alison G. (London School of Economics) **0.8** (0.6)

An automatic method of solving discrete programming problems—*In English*

Econometrica (1960) **28**, 497-520 (8 references, 9 figures)

A generalisation of the classical linear programming problem would be to relax the continuity condition on the non-negative variables which are subject to a system of linear inequalities. This paper presents a numerical algorithm for the solution of programming problems where all or some of the variables are discrete. In some problems the discrete variable constraints are both significant and costly. The algorithm in the paper is simple and lends itself to automatic computations. An alternative approach was given by Gomory [*Bull. Amer. Math. Soc.* (1958) **64**, 275-278] and extended by Beale [*SRG Princeton* (1958) Tech. Rep. 19].

After sections on the formulation of the problem and the description of the method a numerical example is given. Two computational routines, one based upon the solution of many simple linear programmes and the other on a parametric linear programming, are given in Appendix I. In Appendix II a solution is given to the non-negative integer problem originally stated in a paper by Markowitz & Manne [*Econometrica* (1957) **25**, 84-110].

The main method of the paper is directed towards overcoming the difficulty that because a set of feasible solutions to the discrete variable problems is not convex, most methods demonstrate only the existence of a local optimum. The method proposed here is to make systematic parallel shifts in the functional hyperplane until a print within the ordinary linear programming set is found which has integral co-ordinates in the specified dimensions. Although this is obvious in principle, the difficulty of visualising a hyperplane in n-dimensions with a point whose x co-ordinates are all integers, makes it necessary to devise rules to switch attention between possible points so as not to ignore any integer point.

(W. R. Buckland)

2/610

TINTNER, G. (Iowa State University, Ames) **0.8** (0.6)
A note on stochastic linear programming—*In English*
Econometrica (1960) **28**, 490-495 (15 references, 3 tables)

An ordinary linear programming problem has the system of inequalities, the availabilities and coefficients of the objective function as assumed or given quantities. If these parameters become random variables in order to improve the realisation of the model, then the problem becomes one of stochastic linear programming.

A passive approach to this problem is to approximate the distribution of the objective function and to let the decisions be based thereon. The active approach is to let the decision variables be the amount of resources devoted to the various activities within the scope of the programme. This note investigates the numerical methods for optimising the various distributions of resources. Three examples are given.

(W. R. Buckland)

2/611

VOTAW, D. F., Jr. (Yale University) **0.8** (-.-)
Statistical programming—*In English*
Ann. Math. Statist. (1960) **31**, 1077-1083 (8 references)

The author states, "A 'statistical programming' problem is encountered when the information about one or more constants in a programming problem is statistical." Two programming problems are mentioned, firstly, a special case of the classical Hitchcock-Koopmans transportation problem; and secondly, the personnel assignment problem. The first problem is to find an $n \times n$ matrix (x_{ij}) of real numbers such that

(i) $\sum_{i,\, j=1}^{n} c_{ij}x_{ij}$ is minimised, subject to the constraints $\sum_{i=1}^{n} x_{ij} = 1;\ j = 1, 2, \ldots, n.$

(ii) $\sum_{j=1}^{n} x_{ij} = 1;\ i = 1, 2, \ldots, n.$ $x_{ij} \geqq 0$ where the matrix, (c_{ij}), is given beforehand. The assignment problem is similar, except that (i) is to be maximised, and the nonlinear constraint

(iii) each $x_{ij} = 0$ or 1 is introduced.

Of the two, the assignment problem is discussed in detail. In this case (iii) dictates that there are $n!$ admissible matrices (x_{ij}). The problem is to find a permutation $(j'_1, \ldots, j'_n)$ such that

(iv) $c_{1j'_1} + \ldots + c_{nj'_n} = \max\limits_{(j_1,\, \ldots,\, j_n)} (c_{1j_1} + \ldots + c_{nj_n})$ where $(j_1, \ldots, j_n)$ denotes any permutation of $(1, \ldots, n)$. The sum $(c_{1j_1} + \ldots + c_{nj_n})$ is termed a "permutation sum".

The situation treated in this paper arises if the values of the c_{ij}'s are not all known but information regarding them is available. The notion of "programming by estimation" is introduced. Estimates of the $n!$ permutation sums based on observed values are considered and the permutation corresponding to the largest such estimate is selected.

Programming by estimation is compared with purely random programming, in which a uniform probability $1/n!$ is assigned to each permutation. A theorem is stated and proved to the effect that under rather general conditions, programming by estimation is uniformly at least as good as purely random programming. Under slightly more restrictive conditions, programming by estimation is uniformly better. An example where random programming is better is given. A more general form of the assignment problem is also briefly discussed.

(R. M. Durstine)

2/612

BARTON, D. E., DAVID, Florence N. & FIX, Evelyn (Univ. Coll., London and Univ. Cal., Berkeley) **1.1** (0.4)

The polykays of the natural numbers—*In English*
Biometrika (1960) **47**, 53-60 (7 references, 1 table)

The single-suffix polykays of the finite population F_N, of numbers 1, 2, ..., N are obtained explicitly in terms of Bernoulli numbers and also the polynomials of argument $(N+1)$. This was done by using Aty's formulae, for which a general proof has been given by Barton & David [*Biometrika* (1961) **48**, 190-191], for the moments of the sample mean of n elements taken from a finite population in conjunction with the moment generating function of that mean when the population is also finite (Wilcoxon's statistic). This is achieved by using a limiting process equivalent to " equating coefficients " in a polynomial identity.

The unitary polykays of F_N (with suffix 1^r) are shown directly to be simple multiples of Nörlund's generalised Bernoulli numbers of order $N+1$. A series of relations giving polykays of suffixes 2^s1^r, 2^s, 32^s1^r, 3^r2^r, 42^r in terms of the unitary polykays for general finite populations are derived as also are recurrence relations for polykays: for example between those of suffices 3^r, 53^{r-1}, 4^23^{r-2} and 3^r2 and between those of suffices rs and rl^2. These are used to build up tables of the polykays of F_N to weight eight.

The use of these tables is exemplified by obtaining the first four central moments of Mood's rank-dispersion statistic; that is a polynomial in k_2 and k_1 in samples from F_N. These are compared with the same moments derived by the use of Aty's formulae, noting that Mood's statistic may also be treated as a sample mean from a finite population of squares of natural numbers. Some polykays are given for this latter population.

(D. E. Barton)

2/613

BAXTER, G. & SHAPIRO, J. M. (Univ. Minnesota, Minneapolis and Ohio State Univ.) **1.0** (10.1)

On bounded infinitely divisible random variables—*In English*
Sankhyā (1960) **22**, 253-260 (2 references)

Let X be an infinitely divisible random variable. Let $M(u)$ and $N(u)$ denote the non-decreasing functions determining the representation of the logarithm of the characteristic function of X. In this paper the author proves that for the existence of a constant A such that $\Pr[X>A] = 0$, that is to say that X is bounded from above, it is necessary and sufficient that

(i) $N(u) = 0$ for $u>0$,
(ii) $\sigma^2 = 0$,
(iii) $\lim_{\varepsilon\to 0} \int M(u)\,du = 0$, when the region of integration for u is $-1<u<-\varepsilon$.

A similar result is stated for the case where X is bounded from below. This generalises a well-known result that a bounded infinitely divisible random variable is necessarily a constant. The results are then applied to give information regarding the monotonicity of the sample functions of certain stochastic processes.

(R. Ranga Rao)

2/614

BIRNBAUM, Z. W., ESARY, J. D. & SAUNDERS, S. C. (Univ. Washington and Boeing Scientific Res. Lab., Seattle) **1.1** (10.8)

Multi-component systems and structures and their reliability—*In English*
Technometrics (1961) **1**, 55-77 (4 references, 1 table, 2 figures)

The first part of this paper presents formal properties of general classes of multi-component systems which contain two terminal networks and most other kinds of systems considered previously as special cases. The second part applies the properties to the calculation of reliability where reliability is defined as the probability that a structure will perform the task for which it was designed. This definition implies that a structure can only perform or fail.

Consider a structure of n components: the state of all components is described by a vector $\boldsymbol{x} = (X_1, \ldots, X_n)$ where the random variable $X_i = 1$ if the component performs and $X_i = 0$ if it fails. Assume $\Pr(X_i = 1) = p$ and $\Pr(X_i = 0) = 1-p$ for $i = 1, 2, \ldots, n$. For certain of the 2^n possible vectors the structure will perform and for the rest it will fail. Let $\phi(\boldsymbol{x}) = 1$ if $\boldsymbol{x}$ is a vector for which the structure performs and $\phi(\boldsymbol{x}) = 0$ if $\boldsymbol{x}$ is a vector for which it fails. $\phi(\boldsymbol{x})$ is a random variable called the structure function or, in short, structure. Reliability is then defined as $\Pr\{\phi(\boldsymbol{x}) = 1\} = \mathscr{E}[\phi(\boldsymbol{x})] = h(p)$. Under mild assumptions which are met by most practical situations, $h(p)$ is an S-shaped function. Thus, there exists a critical value, p_0, of p such that for $p > p_0$ the reliability of the system is greater than the reliability of a single component and for $p < p_0$ it is smaller. It is shown that by using replicas of the system instead of single components, a system is obtained with reliability arbitrarily close to one if $p > p_0$; but with reliability arbitrarily close to zero if $p < p_0$.

(A. P. Berens)

2/615

BREIMAN, L. (University of California, Los Angeles) **1.6** (10.1)

The strong law of large numbers for a class of Markoff chains
Ann. Math. Statist. (1960) **31**, 801-803 (2 references)

In this paper the author proves the following result: let $(\Omega, \mathscr{B})$ be a measurable space and let $X_0, X_1, \ldots$ be a Markoff chain with a compact Hausdorf state space Ω and transition probabilities $\Pr(A \mid x)$ with the usual assumption of measurability on Ω for each $A \in \mathscr{B}$. Let $\mathscr{C}$ be the class of continuous real-valued functions on Ω and define an operator T on $\mathscr{C}$ into $\mathscr{C}$ by $T\phi(x) = \mathscr{E}(\phi(X_1) \mid X_0 = x)$. If the Markoff chain satisfies (i) and (ii), where (i) states there exists a unique invariant measure π such that $\pi(A) = \int \Pr(A \mid x)\, \pi(\phi x)$ for all $A \in \mathscr{B}$ and (ii) states if $f \in \mathscr{C}$ then $Tf \in \mathscr{C}$, it follows that for any $\phi \in \mathscr{C}$ and $x \in \Omega$ we have

$$\sum_{n=1}^{n} \phi(X_n)/N \to \mathscr{E}_\pi \phi(X_1)$$

almost surely P_x, as $N \to \infty$ where P_x is the probability measure for $(X_1, X_2, \ldots)$ given $X_0 = x$.

The proof is accomplished through two lemmas, firstly, that for any $\phi \in \mathscr{C}$, $\sum_1^N T^n \phi/N$ converges uniformly to $\mathscr{E}_\pi \phi(X_1)$, and secondly, that

$$(1/N) \sum_k^N [\mathscr{E}(\phi(X_n) \mid X_{n-k+1}) - \mathscr{E}(\phi(X_n) \mid X_{n-k})]$$

tends to zero almost surely for each k. The first, well known in linear space theory, is proved for completeness: the second follows from the Markoff property and a text-book stability theorem for a sequence of random variables centred at conditional expectations given the predecessors.

This result established that for learning models

$$\sum_1^N X_i/N \to \mathscr{E}_\pi X_1$$

almost surely where π is the unique invariant initial distribution.

(S. C. Saunders)

2/616

EHRENFEUCHT, A. & FISZ, M. (Math. Inst., Polish Acad. Sci., Warsaw) **1.5** (1.6)

A necessary and sufficient condition for the validity of the weak law of large numbers—*In English*

Bull. Acad. Polon. Sci. III (1960) **8**, 583-585

The following theorem is proved: Let $\{x_i\}$; $i = 1, 2, \ldots$, be a sequence of independent and equally distributed random variables, and let $\phi(t)$ denote their common characteristic function. The relation

$$\lim_{n\to\infty} \Pr\left[\,\left|\,(x_1+\ldots+x_n)/n-a\,\right|>\varepsilon\right]=0$$

holds for any $\varepsilon>0$ if and only if $\phi(t)$ is derivable at $t = 0$ and $\phi'(0) = ai$.

(L. Kubik)

2/617

ERDÖS, P. & RÉNYI, A. (Math. Inst., Hungarian Academy of Sciences, Budapest) **1.3** (1.5)

On the strength of connectedness of a random graph—*In English*

Acta Math. Acad. Sci. Hung. (1961) **12**, 261-267 (5 references)

The following problem is posed by the authors: let there be given n labelled points $V_1, V_2, \ldots, V_n$. Let us choose at random N edges among the $\binom{n}{2}$ possible edges connecting these n points so that each of the $\binom{\binom{n}{2}}{N}$ possible choices chould be equiprobable. Denote by Γ_{nN} the random graph thus obtained. Let $d(V_k)$ denote the valency of a point V_k, the number of edges starting from V_k, and, for a graph G, $c(G) = \min\limits_{1\leq k\leq n} d(V_k)$. Let $c_p(G)$ and $c_v(G)$ denote the least numbers k and l such that deleting k appropriately chosen vertices, together with all edges starting from these vertices or l edges from G the resulting graph is not connected. Let $v_r(G)$ denote the number of vertices of G which have the valency r; $r = 0, 1, 2, \ldots$.

The authors prove that if

$$N(n) = \tfrac{1}{2}n\log n+\tfrac{1}{2}r\,n\log\log n+\alpha n+o(n)$$

where α is a real constant and r a non-negative integer then (Theorem 2)

$$\lim_{n\to+\infty} \Pr\left[c_p(\Gamma_{n,\,N(n)})=r\right]=\lim_{n\to\infty}\Pr\left[c_e(\Gamma_{n,\,N(n)})=r\right]$$
$$=\lim_{n\to+\infty}\Pr\left[c(\Gamma_{n,\,N(n)})=r\right]=1-\exp\left(-e^{-2\alpha}/r!\right)$$

and the distribution of $v_r(\Gamma_{n,\,N(n)})$ tends to the Poisson distribution with expectation $e^{-2\alpha}/r!$ (Theorem 3).

Theorem 2 is a generalisation of a previous result, see Erdös & Rényi, "On random graphs. I." [*Publ. Math. Debrecen* (1959) **6**, 290-297: abstracted in this journal No. **2**/238, **1.3**].

(K. Sarkadi)

2/618

GHELLINCK, G. de (Operational Research Centre, Brussels) **1.8** (0.8)
Sequential decision problems—*In French*
Cahiers Cent. Études Recherche Opérat. (1960) **2**

This paper is concerned with multi-stage decision problems. In such a problem a decision concerning a system is to be made from a set of alternatives at successive intervals of time. The outcome of any decision is to define a distribution of transition probabilities as well as an expected reward during the period running until the next decision.

The problem is to find a decision policy which in the long run yields the maximum expected reward per period. The number of states of the system as well as the number of alternatives available in any state are both assumed to be finite. There are three different methods to solve this kind of problem; the purpose of the paper is to analyse their mutual relations.

The dynamic programming approach as presented by Bellman, see " Functional approximations and dynamic programming " [*Math. Tab., Wash.* (1959) **13**, 247-251: abstracted in this journal No. **1**/523, **11.5**], is the more general case. This consists in assuming that the process ends at a given time. Using recurrently the " principle of optimality " one finds for each state of the system and for each period an optimal decision as well as the corresponding maximum expected reward: such a procedure converges to a permanent optimal policy.

In contrast with this " value-iteration " approach, Howard recently presented a " policy-iteration " method that uses the finite features of the model more efficiently and is focussed on permanent policies only. Starting from an arbitrary decision policy, the algorithm first determines relative " values " associated with every state of the system. These values then form a criterion that leads toward an improved policy, if any. After a finite number of such steps, one reaches the optimal policy.

As a third method, one may state the problem in terms of linear programming. The non-negative variables are probabilities imposed upon all the available alternatives as a mixed strategy. These probabilities are constrained by linear equalities in order to be consistent with steady-state decisions. The expected reward per period is the linear objective function.

After a brief summary of these three methods, the

continued

GHELLINCK, G. de (Operational Research Centre, Brussels) **1.8** (0.8)
Sequential decision problems—*In French*
Cahiers Cent. Études Recherche Opérat. (1960) **2**

continued

paper analyses their common features and states several propositions. As an illustrative example a replacement problem is successively handled by the three methods.

Whatever was said about the processes that make changes at discrete intervals of time, may be adapted to continuous-time processes. The critical feature to be maintained, is the finiteness of both the number of states and the number of available alternatives.

(G. de Ghellinck)

GŁADYSZ, S. (Math. Inst., Polish Academy of Sciences, Wrocław) **1.5** (10.0)

An ergodic paradox—*In German*

Coll. Mat. (1960) 7, 245-249

Let X be a space, $\mathcal{E}$ a class of its subsets E, and $m(E)$ a probability measure defined on $\mathcal{E}$. The author considers a measurable measure-preserving transformation Tx and a " distance " $\rho(x, y)$ defined on X and such that $\rho(x, Tx)$ is a measurable function with finite expectation. In a sequence of transforms $x, Tx, T^2x, \dots$ of a given point x we can distinguish those points T^nx which are in a given set E. If n_1 and n_2 are two consecutive integers for which it is so, we call the sum

$$\rho(T^{n_1}x, T^{n_1+1}x) + \rho(T^{n_1+1}x, T^{n_1+2}x)+\dots+\rho(T^{n_2+1}x, T^{n_2}x)$$

the length of way made by x between these two consecutive sojourns in E. This gives rise to a sequence $r_1(x), r_2(x), \dots$, of lengths-of-ways made by x between consecutive sojourns in E. Under usual conditions on T it can be shown with the aid of Birkhoff's ergodic theorem that for a given E with $m(E)>0$ the mean of r's, that is

$$\lim_{k\to\infty} \frac{1}{k}[r_1(x)+\dots+r_k(x)],$$

exists with probability one and does not depend on x.

However, as shown by the author, it cannot be so for the mean of every sth length. Namely; an example is given where for $s = 2$ the mean

$$\lim_{k\to\infty} [r_2(x)+r_4(x)+\dots+r_{2k}(x)]/k$$

exists with probability one, but takes on essentially two different values.

(S. Zubrzycki)

2/621

GYIRES, B. (University of Debrecen, Hungary) **1.5** (0.6)

A generalization of the local form of the central limit distribution theorem—*In Hungarian*

Magy. Tud. Akad. III. Oszt. Közl. (1960) **10**, 469-479 (3 references)

The author generalises the local form of the limit distribution theorem concerning a sequence of equally distributed independent random variables. He starts by observing that the local limit distribution theorem is equivalent to the asymptotical investigation of the powers of a stochastic matrix, and is infinite in four directions. The author investigates a stochastic matrix infinite in four directions, which is generalised Toeplitzian in the sense that the elements of the matrix are replaced by quadratic matrices of finite order. It is shown that, if the elements of the nth power of such a matrix are denoted by $A_k^{(n)}$; $k=0, \pm 1, \pm 2, \dots$, then, under certain conditions and with suitably chosen constants B_n, the relation

$$B_n A_k^{(n)} - \frac{1}{\sqrt{2\pi}} e^{\frac{z^2}{2}} \boldsymbol{P} \to 0$$

holds for $n\to\infty$, where $\boldsymbol{P}$ is a stochastic projector matrix and the convergence is uniform in k.

A probabilistic interpretation of this result is also given. The results of the paper are obtained by the use of a generalisation of the method of characteristic functions.

(T. Balogh)

2/622

KARLIN, S. & PROSCHAN, F. (Stanford University) **1.0** (2.0)

Pólya type distributions of convolutions—*In English*

Ann. Math. Statist. (1960) **31**, 721-736 (16 references)

A function $f(x, y)$ of two real variables is said to be totally positive of order k (TP_k) if for all $1 \leqq m \leqq k$, all $x_1 < x_2 < \ldots < x_m$ and $y_1 < y_2 < \ldots < y_m$ the determinant $|f(x_i, y_j)| \geqq 0$.
If a TP_k function $f(x, y)$ is a probability density in x with respect to a σ-finite measure $\mu(x)$ for each fixed y, then $f(x, y)$ is said to be of Pólya type of order k (PT_k). If a PT_k frequency $f(x, y)$ can be written in the form $f(x-y)$ then $f(u)$ is called a Pólya frequency density of order k (PF_k).

The author studies these classes of functions and obtains interesting results concerning the n-fold convolutions of PF_k's. He also studies the variation-diminishing and smoothing properties of TP_k functions as well as the properties of compound distributions composed from PT_k densities.

(E. Lukacs)

2/623

KRAFT, C. A., PRATT, J. W. & SEIDENBERG, A. (Michigan State Univ., Univ. Chicago & Harvard Univ., and Univ. California, Berkeley) **1.1** (1.0)

Intuitive probability on finite sets—*In English*

Ann. Math. Statist. (1959) **30**, 408-419 (4 references)

This paper presents further results on the axiomatic development of intuitive probability on finite sets introduced by de Finetti [see L. J. Savage, *Foundations of Statistics* (1954) New York: Wiley]. If one presumes that, for every pair of compound events, one can judge which is more likely or whether they are equally likely, then it is possible to assign a numerical probability to each elementary event in a way consistent with these judgments; that is, does there exist a strictly, or almost strictly, agreeing measure? De Finetti conjectured " yes."

The authors show that the answer is " no " by presenting counter examples, in which there are five elementary events. They develop a necessary and sufficient condition that an ordering of the compound events which satisfies de Finetti's axioms does arise from a measure. By artificially enlarging the set of elementary events, based on a construction in which each event is decomposed into two equally likely and exclusive events, and adding intuitively justifiable axioms, they are able to prove in this new system that any ordering is compatible with a measure.

(W. J. Hall)

2/624

ONICESCU, O. (University " C. I. Parhon ", Bucharest) **1.1** (1.0)
Sum functions on a Boolean algebra—*In French*
Bull. Math. Soc. Sci. Math. Phys. R.P.R. (1959) **3**, 77-91

A new foundation of probability theory is given, replacing random variables by sum functions. Thus, let (Ω, P) be a probabilised Boolean σ-algebra. By definition, a sum function on Ω is a real valued completely additive function F on Ω; a P-sum function on Ω is a sum function F absolutely continuous with respect to P. A P-sum function is simple if it is of the form $F(\omega) = \sum_{j=1}^{n} a_j P(\theta_j \cap \omega)$ where a_j, $1 \leqq j \leqq n$, are given real numbers and θ_j, $1 \leqq j \leqq n$, is a finite partition of the unit element of Ω.

Operations with simple P-sum functions as additivity, multiplication and convolution are defined; moreover, a theorem concerning unicity of the representation of simple P-sum functions is proved. It is also shown that every P-sum function F of the form $G-H$, $G(\omega) \geqq 0$, $H(\omega) \geqq 0$, $\omega \in \Omega$, is the uniform limit of a sequence of simple P-sum functions. Further, a canonical representation of simple P-sum functions is provided and its properties are investigated.

Finally, two theorems concerning the local and integral representation of P-sum functions are given and operations with P-sum functions in the wide sense are defined.

(R. Theodorescu)

2/625

PALÁSTI, Ilona. (Math. Inst., Hungarian Academy of Sciences, Budapest) **1.4** (11.7)
On some random space filling problems—*In English*
Publ. Math. Inst. Hung. Acad. Sci. (1960) **5**, 353-360 (3 references, 1 table, 1 figure)

The following two-dimensional model of random space filling is considered by the author in this paper. Domains which are congruent and parallel with a given domain D of unit area, are placed at random into a rectangle with sides $x \gg 1$, $y \gg 1$. If a domain intersects any of previously placed ones it is discarded. The random space filling is finished when there will be no further possibility of placing a domain without intersection. Let $M(x, y)$ denote the mean value of the number of domains placed by applying this model.

The author proves that the limit

$$\lim_{\substack{x \to \infty \\ y \to \infty}} \frac{M(x, y)}{xy} = \alpha(D)$$

exists if the following hypothesis is true: there exists a constant $A > 0$ depending on D, such that for any $x_1 > 0$, $x_2 > 0$ and $y > 0$ we have

$$|M(x_1+x_2, y) - M(x_1, y) - M(x_2, y)| \leqq Ay$$

respectively for any $x > 0$, $y_1 > 0$, $y_2 > 0$

$$|M(x, y_1+y_2) - M(x, y_1) - M(x, y_2)| \leqq Ax.$$

The proof is based on a lemma which is a generalisation of a theorem of Hyers, related to functional inequalities.

The author deals with the special case when D is a rectangle. An analogous problem in the one-dimensional case was solved by Rényi [*Publ. Math. Inst. Hung. Acad. Sci.* (1958) **3**, 109-127: abstracted in this journal No. **1**/18, **1.4**], and the limiting relation

$$\lim_{x \to \infty} M(x)/x = C \sim 0{\cdot}748$$

was found. The author's conjecture $\alpha(D) = C^2$ is in very good agreement with the results of some Monte Carlo experiments carried out.

The corresponding n-dimensional ($n \geqq 3$) random space filling problems are said to be treated in the same way as the two-dimensional one.

(G. Bánkövi)

2/626

SHAPIRO, J. M. (Ohio State University, Columbus, Ohio) **1.5** (2.9)
Sums of small powers of independent random variables—*In English*
Ann. Math. Statist. (1960) **31**, 222-224 (3 references)

Given a double sequence of infinitesimal independent random variables, $\{x_{nk}\}$, the author studies the distribution function $F_n^r(x)$ of

$$S_n^r = \sum_k | x_{nk} |^r - B_n(r);$$

where $0<r<1$ and $B_n(r)$ is a suitably chosen constant. Necessary and sufficient conditions are given for $F_n^r(x)$ to converge ($n \to \infty$) to a distribution function [$F_r(x)$]. If $F_r(x)$ converges ($r \to 0+$) to $H(x)$, $H(x)$ is the distribution function of the sum of two independent random variables, one Poisson and the other Gaussian.

In a previous paper [*Ann. Math. Statist.* (1958) **29**, 515-522] he considered the case of $r \geqq 1$.

(R. L. Anderson)

2/627

TROTTER, H. F. (Kingston, Canada) **1.5** (-.-)
An elementary proof of the central limit theorem—*In English*
Arch. Math., Karlsruhe (1959) **10**, 226-234 (2 references)

The proof given by the author is similar to that given by Lindeberg [*Math. Zeit.* (1922) **15**, 211-225] and is elementary in the sense that it avoids the use of characteristic functions. The similarity ends when the author introduces the new idea of the characterisation of the distribution of a random variable X by a contracting operator T_x in the space of continuous functions $(T_X f)(\xi) = \mathscr{E}\,[f(X+\xi)]$.
This operator has the property $T_{X+Y} = T_X T_Y$ for the independent random variables, X and Y. For convergence in distribution it is sufficient to have

$$(\mathrm{T}_{X_v} - T_Y) f(0) \to 0$$

for twice differentiable functions f.

This criterion is applied by means of Taylor's expansion and the result is generalised to several dimensions.

(D. Morgenstern)

2/628

ULLRICH, M. & URBANIK, K. (Inst. Radio Eng. and Electronics, Czech. Acad. Sci.,
Prague and Math. Inst., Polish Acad. Sci., Wrocław) **1.5** (1.0)
A limit theorem for random variables in compact topological groups—*In English*
Coll. Mat. (1960) **7**, 191-198

Let X_1, X_2, ..., be a sequence of independent random variables with values in a compact topological group G, which are symmetrical. The authors consider probability distributions of random variables $Y_n = X_1 \cdot X_2 \cdot \ldots \cdot X_n$, where the product is taken in the sense of group multiplication. The question considered is under what conditions imposed on the probability distributions, the μ_i of the X_i's, the probability distributions ν_n of the Y_n's are weakly convergent to the uniform probability distribution over G. The authors prove that the existence of a positive probability distribution λ over G, that is such that $\lambda(V)>0$ for every non-empty open subset V of G, and of a sequence $a_1, a_2, \ldots$ of numbers with $0 \leqq a_n \leqq 1$ such that for every Borel subset E of G we have

(i) $\mu_n(E) \geqq a_n \lambda(E)$

and

(ii) $a_1+a_2+\ldots = \infty$

is a sufficient condition for this. Moreover the authors show that the condition (ii) is essential in the sense that if only the series in it is convergent, then there is a sequence of X_i's so that the probability distributions of Y_n's do not converge to the uniform probability distribution over G.

(S. Zubrzycki)

2/629

WOLFF, P. M. de. (Philips Labs., Irvington-on-Hudson, N.Y.) **1.2** (1.4)
Particle statistics in X-ray diffractometry. I. General Theory—*In English*
Appl. Sci. Res., Hague (1958) **7**, 102-112 (1 reference, 3 figures)

In X-ray diffraction by powders, a statistical error in the measured intensity results from the finite number of crystals contributing to it. A recent device for reducing this error, the rotating sample holder or specimen spinner, is not covered by existing formulae for stationary specimens. The present paper establishes an expression for this new situation, which has also been investigated experimentally in the second part of this paper [*J. Appl. Phys*, (1959) **30**, 63-69].

The contribution of each crystallite can be written as a product of functions of mutually independent variables; its volume, and its position and orientation coordinates. The resulting root mean-square error is expressed in the relative standard deviations of those functions. The latter are calculated using somewhat simplified physical assumptions. This is done first for the stationary specimen, where it is found that the form of the distribution of intensity contributions has surprisingly little effect on the statistical error. The same conclusion is reached for rotating specimens, where a very rough estimate of the number of contributing crystals leads to an expression differing only in the numerical factor, 4 against 6·5, from the final expression which is obtained by detailed analysis of the rather intricate geometry of this case.

(P. M. de Wolff)

2/630

EPSTEIN, B. (Wayne State University and Stanford University) **2.6** (1.5)
Elements of the theory of extreme values—*In English*
Technometrics (1960) **2**, 27-41 (8 references)

The purpose of this paper is to describe briefly the basic elements of the theory of extreme values. Frequent examples are given to illustrate both the derivations and applications of the theory.

Throughout the article, the author deals with samples of size n, drawn independently and at random from a parent population with density function $f(x)$ or distribution function $F(x)$. Of particular interest here is the minimum value, y_n, and the maximum value, z_n, from such a sample.

The density function and distribution function for y_n are derived in terms of $F(x)$. Examples are given for parent uniform and exponential populations.

Using a technique of Cramér, the author asserts that the sequence of random variables, y_n, converges in distribution to a random variable y, whose distribution is more easily evaluated. Examples are given where the parent populations are selected to illustrate Gumbel's types 1, 2 and 3 asymptotic distributions of smallest values. The conditions for each type of asymptotic distribution is given.

The equivalent results for the largest value, z_n, are also derived, with corresponding examples and their relationship to Gumbel's classification of asymptotic distributions of largest values.

For the case when $f(x)$ is differentiable, the author shows several methods by which the modes of the densities of y_n and z_n may be found. These methods are carried out for some of the preceding examples.

The article concludes with remarks on the applications of this theory to a variety of fields.

(L. A. Foster)

2/631

FREIMER, M. & GOLD, B. (Lincoln Lab., Massachusetts Inst. Tech.) **2.1** (2.9)
Note on the distribution of locally maximal elements in a random sample—*In English*
Ann. Math. Statist. (1960) **31**, 518 (3 references)

This note presents a simpler second factorial moment of the distribution referred to in the title than presented by Glasgow [*Ann. Math. Statist.* (1959) **30**, 586-590; abstracted in this journal No. **1**/190, **2.1**].

(R. L. Anderson)

2/632

GOLOVIN, N. E. (Natl. Aeronautics & Space Admin., Washington, D.C.) **2.5** (1.0)
An approach to a reliability prediction programme—*In English*
Annual Conv. Trans., Amer. Soc. Qual. Contr. (1960) 173-182 (4 references, 5 figures)

Before there is useful operating experience for a system, reliability prediction is valuable in forecasting operational feasibility and costs versus reliability. A failure is defined as occurring at the time any relevant physical characteristic of the system attains a value outside of a specified range. A full reliability analysis would distinguish between catastrophic and degradation failures. The procedure recommended here is to start with system failures and work back through sub-systems and components to the data on parts failures. A given set of mathematical relationships can be taken which allow calculation of system outputs from sub-system outputs. In this way, system failures can be traced to sub-system failures. The effects of variation in sub-system outputs on the system output can be studied. Grouping sub-systems into independent groups allows the use of the product rule for determining probabilities. Components of a sub-system and parts of a component can be treated similarly. Any prediction of reliability must be done for a specified set of environmental conditions.

The reliability, that is the probability that a part survives to time t, can be expressed in general as

$$R_p(t) = \exp\left[-\int_0^t \lambda_p(t)\,dt\right]$$

where $\lambda_p(t)$ is the failure rate of the pth part per unit of time. Frequently, $\lambda_p(t)$ increases rapidly at first during an " infant mortality " period, decreases to a fairly constant level, chance failures, during a " normal operating " period, and finally increases during a " wear-out " period. For a component with n independent parts

$$R_c(t) = \exp\left[-\int_0^t \sum_{i=1}^{n} \lambda_i(t)\,dt\right].$$

If there are k identical parts in a redundant arrangement, the reliability of the group is

$$R_g(t) = 1-[1-R_p(t)]^k.$$

The validity of a reliability prediction depends on the ability to identify the independent groups of a system and on the information available for the failure rates of the groups, $\lambda_g(t)$.

2/633 (D. W. Gaylor)

JOHN, S. (Indian Statistical Institute, Calcutta) **2.9** (0.3)
On the evaluation of the probability integral of a multivariate normal distribution—*In English*
Sankhyā (1959) **21**, 367-370 (7 references)

It is shown by the author in this paper that the probability integral of a multivariate normal distribution of order k can be expressed as the sum of k terms each of which is the product of a univariate normal density and multivariate normal integral of order $k-1$. To illustrate the application of the formula an example is given.

(T. V. Hanumantha Rao)

2/634

KOTLARSKI, I. (Warsaw) **2.9** (0.4)

On random variables whose quotient follows the Cauchy law—*In English*

Coll. Mat. (1960) **7**, 277-284

The author considers a class $\mathscr{X}$ of random variables ξ with the property that the quotient of two independent random variables having the same distribution as ξ has Cauchy's distribution. Owing to the connections between Mellin transforms $h(z) = E\,|\,\xi\,|^z$, z complex, and characteristic functions $\phi(t) = \mathscr{E}e^{it\xi}$, where t is real of random variables ξ, he characterises the class $\mathscr{X}$ in terms of Mellin transforms.

(S. Zubrzycki)

2/635

PATIL, G. P. (Indian Statist. Inst. and Univ. Michigan, Ann Arbor) **2.5** (0.4)

On the evaluation of the negative binomial distribution with examples—*In English*

Technometrics (1960) **2**, 501-506 (11 references)

In this paper the author proposes an evaluation for

$$Y(r,\, p,\, k) = \sum_{x=0}^{r} y(x,\, p,\, k)$$

of the negative binomial density function

$$y(x,\, p,\, k) = \binom{k+x-1}{x} p^k q^x,\; p+q = 1.$$

If $k = 1, 2, \ldots$, then

$$Y(r,\, p,\, k) = \frac{1}{\beta(k, r+1)} \int_0^p u^{k-1} v^r\, du,$$

$u+v = 1$; namely the incomplete beta function.

A short proof of this is given by the author and is as follows:

$Y(r$, p, k$)$ = Pr [at most $k+r$ independent trials are required to achieve k successes when p is the probability of any single success]

= Pr [at least k successes occur in $k+r$ trials]

$$= \sum_{x=k}^{k+r} \binom{k+r}{x} p^x q^{k+r-x},$$ which by a well-known theorem connecting the binomial and beta distribution

$$= I_p(k,\, r+1) = \frac{1}{\beta(k,\, r+1)} \int_0^p u^{k-1} v^r\, du.$$

A numerical example is given.

(E. H. Lehman, Jr.)

2/636

MENDENHALL, W. & LEHMAN, E. H., Jr. (North Carolina State College) **2.4** (4.3)
An approximation to the negative moments of the positive binomial useful in life testing—*In English*
Technometrics (1960) **2**, 227-242 (10 references, 4 tables, 1 figure)

The Weibull density function

$$f(t;\ m,\alpha) = (m/\alpha)\, t^{m-1} \exp\left(-t^{m/\alpha}\right)$$

is useful in life testing. Assuming m known, n items are placed on test and the test is stopped when time T has elapsed. It is assumed that T is determined in such a way that the probability of zero failures is negligible.

Then the maximum likelihood estimator for α is

$$\hat{\alpha} = \frac{1}{r}\sum_{i=1}^{r} t_1^m + [(n/r)-1]\, T^m,$$

where t_i is the time until the ith ordered failure and r is the number of failures.

The expectation of $\hat{\alpha}$ is shown to be

$$\alpha + T^m\, [n\mathscr{E}(1/r) - (q/p) - 1]$$

where $q = \exp(-T^m/\alpha)$ is the probability of survival until time T, $p = 1-q$, and r is distributed as a positive binomial: $\hat{\alpha}$ therefore is considerably biased.

The variance of $\hat{\alpha}$ then is

$$V(\hat{\alpha}) = T^{2m}\left[V\left(\frac{1}{r}\sum_{1}^{r} x_i^m\right) + n^2 V\left(\frac{1}{r}\right) + 2n \operatorname{cov}\left(\frac{1}{r}\sum x_i^m, \frac{1}{r}\right)\right]$$

The covariance term is shown to be zero and then, after some manipulation,

$$V(\hat{\alpha}) = T^{2m}\left\{n^2 V\left(\frac{1}{r}\right) + \left[\left(\frac{\alpha}{T^m}\right)^2 - \frac{q}{p^2}\right]\mathscr{E}\left(\frac{1}{r}\right)\right\}.$$

Hence, the first two negative moments of r, which is distributed as a positive binomial, are required: to approximate these the " best " beta function is used. It is noted that if y has the beta distribution with parameters $a = (n-1)\,p$ and $b = (n-1)\,q$, then

$$\mathscr{E}\left(\frac{1}{y}\right) = \frac{n-2}{a-1},\quad \mathscr{E}\left(\frac{1}{y^2}\right) = \frac{(n-2)(n-3)}{(a-1)(a-2)}.$$

continued

MENDENHALL, W. & LEHMAN, E. H., Jr. (North Carolina State College) **2.4** (4.3) *continued*
An approximation to the negative moments of the positive binomial useful in life testing—*In English*
Technometrics (1960) **2**, 227-242 (10 references, 4 tables, 1 figure)

Now r can be approximated by ny, for

$$\mathscr{E}(r) = np = \frac{na}{n+b} = \mathscr{E}(ny)$$

and

$$\mathscr{E}(r^2) = npq + n^2p^2 = \frac{n^2 a(a+1)}{(a+b)(a+b+1)} = \mathscr{E}(n^2y^2).$$

Therefore $\mathscr{E}\left(\frac{1}{r}\right)$ and $\mathscr{E}\left(\frac{1}{r^2}\right)$ can be approximated respectively by

$$\mathscr{E}\left(\frac{1}{ny}\right) = \frac{(n-2)}{n(a-1)}$$

and

$$\mathscr{E}\left(\frac{1}{n^2y^2}\right) = \frac{(n-2)(n-3)}{n^2(a-1)(a-2)}.$$

Then $V(1/r)$ is approximated by

$$\frac{(n-2)(n-a-1)}{n^2(a-1)^2(a-2)}.$$

$\mathscr{E}(1/r)$ and $\mathscr{E}(1/r^2)$ are tabulated for $n = 5$, 10, 20, 30, 40, 100 and $p = 0{\cdot}1$, $0{\cdot}3$, $0{\cdot}5$, $0{\cdot}7$, $0{\cdot}9$.

The approximation is good to two significant digits for $np \geqq 5$.

(E. H. Lehman, Jr.)

COHN, R., MOSTELLER, F., PRATT, J. W. & TATSUOKA, M. **3.8** (5.6)
(Rutgers Univ., Harvard Univ., Harvard Univ. and Univ. of Hawaii)
Maximising the probability that adjacent order statistics of samples from several populations form overlapping intervals—*In English*
Ann. Math. Statist. (1960) **31**, 1095-1104 (6 references)

The problem treated here arose in a sociological study, but is stated abstractly as follows, " Provide a test of the hypothesis that several samples of the same size are randomly drawn from possibly different populations, against the alternative that the samples are not independently and randomly drawn from distributions ".

The authors consider the interval, in each of the k samples of size n, between the rth and $(r+1)$st order-statistics. It is determined whether these k intervals have a non-empty intersection; or, equivalently, whether the least of all the rth order-statistics is greater than or equal to the greatest of all the $(r+1)$st order statistics. In this case, one looks to see with what probability it would occur had the samples been randomly and independently drawn from the same population. If this probability is sufficiently small, one rejects the hypothesis that the samples were randomly and independently drawn from, possibly different, populations, since it is shown that the probability of a non-empty intersection, when the samples are drawn from k univariate continuous cumulative distribution functions on the same real line, is greatest when the populations are identical.

After treating the case $r = 1$ in Sections three and four, the authors prove two general theorems in Section five. If x_{ri} denotes the rth observation, in descending order of magnitude in the ith sample, the theorems are firstly, if samples of size n are drawn from each of k continuous distribution functions $y_i(x)$; $i = 1, 2, \ldots, k$, the probability that $\min\limits_i x_{ri} > \max\limits_i x_{r+1,\,i}$ for a given $r\,(= 1, 2, \ldots, n-1)$ is maximised if and only if $y_1 = y_2 = \ldots = y_k$, and secondly under the same conditions, the probability that

$$\min_i x_{ri} > \max_i x_{r+1,\,i} \quad \text{and} \quad \min_i x_{si} > \max_i x_{s+1,\,i}$$

for a given $r\,(= 1, 2, \ldots, n-1)$ and $s\,(= r, r+1, \ldots, n-1)$, is maximised if and only if $y_1 = y_2 = \ldots = y_k$. In each theorem a combinatorial formula for the maximum probability is given. The results extend to three or more corresponding intervals. The authors describe obstacles in the path of extending the results from the univariate to the multivariate case: and remark on some other possible extensions.

2/639 (R. Hooke)

GUPTA, S. S. (Bell Telephone Laboratories, Allentown, Pennsylvania) **3.8** (2.7)
Order statistics from the gamma distribution—*In English*
Technometrics (1960) **2**, 243-262 (7 references, 3 tables)

This paper is concerned with order statistics from the gamma or Pearson Type III distribution. The first four moments about the origin and the mean are calculated and tabled for the kth order statistic from a sample of size N. An expression is derived which would enable one to evaluate the covariance between the mth and nth order statistics.

The distribution function of the kth order statistic is obtained in terms of the incomplete beta function and also in terms of the incomplete gamma function and the cumulative Poisson distribution. An iterative procedure is presented for finding the mode of the kth order statistic and a table is presented evaluating this mode with $k = 1$ and $k = N$ for several values of the gamma-distribution parameter r. If $B_\alpha(k, N-k+1)$ denotes the α-percentage point of the beta distribution, then Y_α, the α-percentage point of the kth order statistic, is the solution of

$$G_r(Y_\alpha) = B_\alpha(k, N-k+1).$$

An iterative procedure was used to solve this equation and Y_α is tabled for certain values of N, k, r and α.

If the gamma density function is considered in the form

$$g_r(x, \theta) = \frac{e^{-x/\theta}\, x^{r-1}}{\theta^r \Gamma(r)}, \quad x>0,\ \theta>0,$$

where θ is to be estimated, then a method of finding the best linear unbiased estimate of θ is presented in the case of censored sampling.

Applications to life tests, extreme values, reliability and maintenance are described and illustrated in some cases.

(A. P. Berens)

2/640

HOGG, R. V. (University of Iowa, Iowa City, U.S.A.) **3.6** (3.5)
Certain uncorrelated statistics—*In English*
J. Amer. Statist. Ass. (1960) **55**, 265-267 (1 reference)

The author proves that an odd location statistic and an even location-free statistic are uncorrelated when sampling from a symmetric distribution. An odd location statistic, T, by definition is transformed similarly to the sample space under (i) translation or (ii) multiplication by -1, while an even location statistic, S, by definition is invariant under (i) or (ii). The theorem of the author is stated as:

Theorem. Let $X_1, X_2, \ldots, X_n$ be a random sample from either a continuous-type or a discrete-type distribution that is symmetric about the point c. If the correlation coefficient of an odd location statistic $T(X_1, X_2, \ldots, X_n)$ and an even location-free statistic $S(X_1, X_2, \ldots, X_n)$ exists, it is equal to zero.

The proof may be outlined as follows: the expectation, $\mathscr{E}[T]$, of T is c. Thus the covariance, V, of T and S is $\mathscr{E}[(T-c)S]$. Employing the transformation $x_i = y_i + c$ and the definitions, it follows that $V = \mathscr{E}[TS]$. By the transformation $y_i = -z_i$ and the definitions it follows that $V = -\mathscr{E}[TS]$. It is thus implied that $V = 0$ and hence the correlation is zero.

The author notes that the theorem is actually valid for every random vector whose distribution is symmetric about $(c, c, \ldots, c)$ in n-dimensional space. Restriction to stochastically independent and identically distributed random variables and to either continuous-type or discrete-type distribution is unnecessary.

(S. Krane)

2/641

RIDER, P. R. (Wright Air Development Centre, Dayton, Ohio) **3.8** (2.9)
Variance of the median of samples from a Cauchy distribution—*In English*
J. Amer. Statist. Ass. (1960) **55**, 322-323 (2 references, 1 table)

The author obtains the integral expressing the exact variance of the median for samples of size $2k+1$ from the Cauchy distribution with location parameter θ. The variance is computed by numerical integration for $k = 2\,(1)\,15$. Corresponding values of the approximate variance formula for large samples are computed for comparison. Values of the intrinsic accuracy, that is the variance of logarithmic derivative of density function, are also given and the three functions are displayed in tabular form.

The large-sample approximation is found to be poor for small samples. For $k = 2$, for example, the exact variance is given as 1·221253, the large-sample approximation as 0·493480 and the reciprocal of the given intrinsic accuracy is 0·572895. For $k = 15$ the corresponding values are 0·087944, 0·079594 and 0·085545.

(S. Krane)

2/642

SARKADI, K. (Hungarian Academy of Sciences, Budapest, Hungary) **3.8** (1.3)
On the median of the distribution of exceedances—*In English*
Ann. Math. Statist. (1960) **31**, 225-226 (4 references)

If we have two random samples of sizes n_1 and n_2 from the same continuous distribution, the number of exceedances (x) is defined by the author as the number of elements of the second sample which surpass at least $n_1 - m + 1$ elements of the first, for a fixed natural number $m \leqq n_1$. The cumulative distribution of x, $W(n_1, m, n_2, x)$ is given; if $n_1 = n_2$, the median is $m-1$.

In this paper it is shown that

$$W(n_1, m_1, n_2, m_2 - 1) + W(n_2, m_2, n_1, m_1 - 1) = 1.$$

(R. L. Anderson)

2/643

SIDDIQUI, M. M. (National Bureau of Standards, Boulder, Col., U.S.A.) **3.8** (4.4)
Distribution of quantiles in samples from a bivariate population—*In English*
J. Res., Nat. Bur. Stand. B (1960) **64**, 145-150 (3 references)

This paper extends the results of former investigators—the exact distribution of quantiles in the univariate case, the distribution of medians in the multivariate case—to the distribution of quantiles in samples from a bivariate population. The general approach follows that of Cramér and Mood. Correlation coefficient between quantiles is estimated by the use of an auxiliary statistic.

The author considers the joint distribution $F(x, y)$ of (X, Y), possessing a probability density function $f(x, y)$, and having $F_1(x)$ and $F_2(y)$ as the respective marginal distribution functions. Let α be a quantile of $F_1(x)$, and β be a quantile of $F_2(y)$. It is assumed that the first and second partial derivatives of $F(x, y)$ are continuous in a neighbourhood of (α, β) and that $f(\alpha, \beta) \neq 0$. A random sample (X_k, Y_k), $k = 1, 2, \ldots, n$, is drawn and the values on each variate are ordered so that $X_i' < X_j'$ and $Y_i' < Y_j'$ if $i < j$. Let i and j be the greatest integers such that $i/n < F_1(\alpha)$, $j/n < F_2(\beta)$, and let M be the number of elements (X, Y) such that $X < X_i'$ and $Y < Y_j'$. The joint distribution of (M, X_i', Y_j') is obtained and is shown to be asymptotically trivariate normal. The parameters of the asymptotic joint distribution of the variates (X', Y') and that of the distribution of Q, where Q is a continuous random variable replacing the discrete random variable M/n such that

$$|Q - M/n| < 1/n,$$

are given in a theorem.

Estimates of α, β, $F(\alpha, \beta)$ and $\rho(\alpha, \beta)$ are given together with examples in setting confidence limits on these parameters.

(H. H. Ku)

2/644

BANERJEE, S. K. (Indian Statistical Institute, Calcutta) **4.4** (5.4)
Approximate confidence interval for linear functions of means of k populations when the population variances are not equal—*In English*
Sankhyā (1960) **22**, 357-358 (2 references)

Let $N_i\,(m_i, \sigma_i^2)$, $i = 1, 2, \ldots, k$, be k normal populations with means m_i and variances σ_i^2. A confidence interval for $\sum_1^k c_i m_i$, where the c_i's are specified constants, is required.

Let $\bar{x}_i$ and s_i^2 be respectively, the sample arithmetic mean and the sample variance computed from a random sample of size n_i from $N_i\,(m_i, \sigma_i^2)$. Let t_i be the upper 50α per cent. point of the distribution of student's t on $n_i - 1$ degrees of freedom. The following inequality is proved:

$$\Pr\left\{\sum_1^k c_i\bar{x}_i - \left(\sum_1^k n_i^{-1} t_i^2 s_i^2\right)^{\frac{1}{2}} \leqq \sum_1^k c_i m_i \leqq \sum_1^k c_i\bar{x}_i + \left(\sum_1^k n_i^{-1} t_i^2 s_i^2\right)^{\frac{1}{2}}\right\} \geqq \alpha$$

This inequality is true whatever may be the values of the parameters m_i and σ_i^2 $(i = 1, 2, \ldots, k)$.

The result obtained by setting $k = 2$, $c_1 = 1$, $c_2 = -1$ in the above expression is of interest in connexion with the problem of testing whether two normal populations, with possibly unequal variances, have equal means.

(S. John)

2/645

BIRNBAUM, A. & HEALY, W. C., Jr. (New York University and Ethyl Corp.) **4.3** (4.7)
Estimates with prescribed variance based on two-stage sampling—*In English*
Ann. Math. Statist. (1960) **31**, 662-676 (17 references, 3 tables)

In most problems of estimation, estimators based on samples of fixed size have precisions which depend on unknown parameters and estimators with prescribed precision are not available without resort to some such method as Stein's two-sample procedure for the estimation of the mean of a normal distribution with unknown variance. For other problems than this the only available general methods are due to Cox and Anscombe but these are based on asymptotic theory and hence are approximate.

Here a new method is given which provides, under conditions satisfied by many common distributions, rules for sampling in two stages so as to obtain an unbiased estimator of a given parameter having variance equal to or not exceeding a prescribed bound. The estimate suggested depends directly only on the observations taken in the second stage; though it depends on them indirectly since the size of the second-stage sample is prescribed by the first-stage sample. At first sight this might suggest a gross waste of information but, in fact, it is shown that the efficiency is so high in a number of cases that the search for more efficient methods is of theoretical rather than practical interest.

The method is applied to estimation of the means of binomial, Poisson and hyper-geometric distributions, scale parameters in general and the gamma distribution in particular, the variance of a normal distribution and a component of variance. Some of these problems are discussed in detail with tabled values of comparisons.

(D. G. Chapman)

2/646

BERKSON, J. (Mayo Clinic, Rochester, Minnesota) **4.1** (6.7)
Problems recently discussed regarding estimating the logistic curve—*In English*
Bull. Int. Statist. Inst. (1960) **37,** III 207-211

This paper discusses criticism made by Silverstone [*J. Amer. Statist. Ass.* (1959) **52**, 567-577] of the recent work by Berkson in which the maximum likelihood estimator for the parameters (α, β) of the logistic function was compared with several other estimators, including the minimum logit χ^2 estimator [*J. Amer. Statist. Ass.* (1953) **48**, 565-599; *J. Amer. Statist. Ass.* (1955) **50**, 130-162; *Proc. Third Berkeley Symposium on Math. Statist. and Prob.* (1956), pp. 1-13]. The chief points of Silverstone's paper were:

1. The maximum likelihood estimator is sufficient, instead of being not sufficient as had been stated.
2. The minimum logit χ^2 estimator is not sufficient in all conditions, though it may be in the conditions explored.
3. The minimum logit χ^2 estimates are not consistent " in Fisher's sense for finite samples," (F.C.) even though they are consistent in the sense of convergence in probability (C.P.).
4. The superior efficiency (mean square error) of the minimum logit χ^2 estimate was reported for certain values of the parameters. However, for values of the parameters " far outside the range " that was examined, " it is doubtful whether such a claim could be upheld."

With regard to these points, the present paper indicates that:

(*a*) With respect to the sufficiency of the maximum likelihood estimate: A necessary and sufficient condition for the estimates $(\hat{\alpha}, \hat{\beta})$ to be sufficient is that no two or more samples having different values for the minimal sufficient statistics $(\Sigma r_i, \Sigma r_i x_i)$ should yield the same values for $(\hat{\alpha}, \hat{\beta})$. There are in fact, samples having different values for $(\Sigma r_i, \Sigma r_i x_i)$ which yield the same estimates $(\hat{\alpha}, \hat{\beta})$; for these samples the estimate is ∞ for either or both of $(\hat{\alpha}, \hat{\beta})$. This establishes that the maximum likelihood estimates $(\hat{\alpha}, \hat{\beta})$ are not sufficient. The proposal of Silverstone of certain degenerate curves as distinguishing " solutions " for the samples concerned, even if correct, is considered irrelevant, for these do not provide distinguishing estimates $(\hat{\alpha}, \hat{\beta})$.

(*b*) With respect to the sufficiency of the minimum

continued

BERKSON, J. (Mayo Clinic, Rochester, Minnesota) **4.1** (6.7) *continued*
Problems recently discussed regarding estimating the logistic curve—*In English*
Bull. Int. Statist. Inst. (1960) **37,** III 207-211

logit χ^2 estimate: In the conditions explored, i.e. $n_i = 10$, $k = 3$, and for an infinite number of other controllable conditions, the estimates are sufficient. It is true, however, that conditions can be defined in which the estimates are not sufficient.

(*c*) With respect to consistency of the minimum logit χ^2 estimate: The estimate is consistent (F.C.) as well as consistent (C.P.). Rao has proved the equivalence of F.C.$\rightleftharpoons$C.P. for conditions which are fulfilled in these experiments [*Bull. Int. Statist. Inst.* (1957) **35**, 25-32]. It is remarked that the maximum likelihood estimate is not always consistent (F.C.) and an example is given.

(*d*) With respect to the relative efficiency of the estimates: It appears to be considered that the minimum logit χ^2 is superior in conditions where the dosages are not exceedingly far from X_{50} but doubt is expressed as regards situations in which all P_i are near unity (or zero). Actually, experiments showed the maximum likelihood estimator to be increasingly inferior to minimum χ^2 estimators as one moved away from the X_{50}. There is a point beyond which comparison cannot properly be made. The maximum likelihood estimator yields infinity as estimates for some samples. The frequency of these samples increases as the disposition of the dosages is removed from the X_{50}, and comparisons were made omitting these samples, but only up to the point where samples insoluble by maximum likelihood are not more than 5 per cent. of the sample population. This is already beyond where practical assays can validly be made. If comparison is to be made for its mathematical statistical interest in regions with all P_i not far from unity (or zero), it will have to be for large sample size where the probability of infinite estimates by maximum likelihood is negligibly small. The author believed that the inferiority of the maximum likelihood estimate found for more centrally placed dosages with small numbers would be confirmed.

(J. Berkson)

BERKSON, J., & ELVEBACK, Lillian (Mayo Clinic, Rochester, and Tulane Univ., New Orleans) **4.1** (2.5)

Competing exponential risks with particular reference to the study of smoking and lung cancer—*In English*

J. Amer. Statist. Ass. (1960) **55**, 415-428 (17 references, 3 tables)

This paper shows that the use of the " mortality ratio " in comparing the numbers of deaths caused by lung cancer in smokers and non-smokers can be very misleading since other causes of death play an important role. The authors utilise the notion of " competing risks " and present a model for the effect of smoking in causing death from a specific disease in which the effect is measured in terms of the net probability of death from this disease.

The probability of death when one risk operates alone is called the net risk or rate. When it operates together with another risk it is called the crude risk or rate. In order to estimate the net risk from the crude risk a simplifying and, in this context, reasonable assumption is made that each net instantaneous risk or force of mortality is constant over the period of observation and that the risks are independent.

For the case in which the numbers of deaths as well as the times when the death occurred are available the authors derive the maximum likelihood estimates, called " time estimates " of the net rates and their variances. For the case in which the numbers of deaths are available but not the times when deaths occurred, the maximum likelihood estimates, called " frequency estimates " and their variances are also provided. The entries in Table 1 show that if the crude death rate Q from all causes is below 0·6 the variances of the above two estimates are close; but for larger values of Q the increase of variance is substantially greater when the information on time of death is not used.

Table 2 gives a list of time estimates and frequency estimates of the parameters involved in a model of exposure to two competing risks such as may be used in the analysis of data giving numbers of deaths among smokers and non-smokers due to lung cancer and to other causes: the formulae for variances are also given. Some data taken from the prospective study sponsored by the American Cancer Society and reported by Hammond and Horn (1958) provides an example. When the net probability of death is used to measure the effect of smoking in respect of a cause of death, smoking is seen to have the greatest effect in increasing the number of deaths from coronary heart disease and the least effect in increasing the number of deaths from lung cancer. If the simple difference of the crude death rates is used, essentially the same conclusions are reached as with the preceding net rates. On the other hand, when the mortality ratio is used, the rate for lung cancer appears very much higher than the rates for other causes of death.

The authors propose the use of net rates as more logical and state that " the finding of associations of smoking with classes of diseases in general, not only, or even chiefly, with lung cancer, is of central importance for an interpretation of the possible biologic significance of these statistical studies ".

(J. Gurland)

continued

BERKSON, J., & ELVEBACK, Lillian (Mayo Clinic, Rochester, and Tulane Univ., New Orleans) **4.1** (2.5) *continued*

Competing exponential risks with particular reference to the study of smoking and lung cancer—*In English*

J. Amer. Statist. Ass. (1960) **55**, 415-428 (17 references, 3 tables)

COHEN, A. C., Jr. (University of Georgia, Athens, USA) **4.3** (2.8)

Estimation in the truncated Poisson distribution when zeros and some ones are missing—*In English*

J. Amer. Statist. Ass. (1960) **55**, 342-348 (9 references, 2 tables, 1 figure)

The author considers the estimation of the parameter of a Poisson distribution for the special case in which all zero observations are missing as well as some of the ones. The maximum likelihood estimating equations for the Poisson parameter and for the proportion of missing ones are obtained. A table and a graphical aid for estimation of the Poisson parameter are included. Asymptotic variances and covariances of estimates are obtained by inversion of the information matrix. A function useful in computing the variance of the Poisson parameter is tabulated.

As an example, the author suitably alters data of Varley on distribution of gall-cells in flower-heads of black knapweed.

(S. Krane)

2/651

COHEN, A. C., Jr. (University of Georgia, Athens, USA) **4.3** (2.8)

Simplified estimators for the normal distributions when samples are singly censored or truncated—*In English*

Technometrics (1959) **1**, 217-237 (14 references, 2 tables, 3 figures)

The author presents a scheme to estimate the mean and variance of a normally distributed population when single truncation or censoring is involved. Only maximum likelihood estimation is used for the cases with restricted sampling in one tail. Other estimating techniques are described from the references along with conditions under which these may be preferable. The most important feature of the paper is that only one simple auxiliary function is required for each of the restricted sample cases covered.

A table and a graph are presented of the single auxiliary estimation function for singly, either right or left, truncated samples. One table and two graphs for varying ratios of the number of censored observations to the total sample size are given of the single auxiliary estimation function for singly, right or left, censored samples. The latter table and graphs are applicable for the cases in which either the point of censoring is fixed or the number of censored observations is fixed. Examples utilising the graphs and tables are presented for the six separate cases of truncated and censored sampling covered in the paper.

In addition, the variances and covariances of the estimates are explained for truncated samples, for fixed point-censored samples and for fixed sample-size censored samples. Examples show how to utilise this information.

(V. L. Anderson)

2/652

CORNISH, E. A. (Division of Math. Statist., CSIRO, Adelaide) **4.8** (3.6)

The simultaneous fiducial distribution of location parameters in a multivariate normal distribution—*In English*

CSIRO Div. Math. Statist. (1961) Tech. Paper No. 8 (8 references, 2 figures)

In this paper the author proves directly the formula for the simultaneous fiducial distribution of the location parameters in a multivariate normal distribution, originally conjectured by Fisher [*J. R. Statist. Soc.* B. (1954) **16**, 212-213].

If $z' = (z_1, \ldots, z_p)$ is distributed in a multivariate normal distribution with mean $\zeta = (\zeta_1, \ldots, \zeta_p)$ and variance-covariance matrix Σ, where this has a non-singular estimate S, based on n degrees of freedom and distributed independently of z, then the simultaneous fiducial distribution of the parameter ζ is of the form

$$\frac{\Gamma\left[\frac{1}{2}(n+\rho)\right]}{(\pi n)^{\frac{1}{2}\rho}\Gamma\left(\frac{1}{2}n\right)}\left\{1+\frac{(\zeta-z)^1 S^{-1}(\zeta-z)}{n}\right\}^{-\frac{1}{2}(n+\rho)}\frac{d\zeta}{|S|^{\frac{1}{2}}}.$$

The author gives details for the fiducial method in practice. Three illustrative examples are given based on rainfall, altitude, latitude and longitude data, pressure and temperature observations at four stations, and shearing strength and density of kauri trees respectively. In conclusion the fiducial method is compared with the method of confidence intervals given by Hotelling [*Ann. Math. Statist.* (1931) **2**, 360-378].

(G. Yeo)

2/653

EPSTEIN, B. (Wayne State University and Stanford University) **4.4** (2.5)

Estimates of bounded relative error for the mean life of an exponential distribution—*In English*

Technometrics (1961) **1**, 107-109 (1 reference, 1 table)

If the underlying distribution of life is exponential, an estimation procedure is presented for mean, θ, having a small relative error. More precisely, an estimate of θ is found which is within a certain percentage, 100δ, of θ with preassigned confidence $1-\alpha$. In life testing this requires observing a sufficiently large number of failures, r.

For the loss function which is zero when $1-\delta \leqq a/\theta \leqq 1+\delta$ and one elsewhere, the minimax estimate a, of θ is found. For small δ, a is approximately equal to $\hat{\theta}$, the maximum likelihood estimate of θ. Then for small δ and small α, the central limit theorem is used to determine how large r must be in order that

$$\Pr\left\{\left|\frac{\hat{\theta}-\theta}{\theta}\right| < \delta\right\} \geqq 1-\alpha.$$

It is also shown how to compute, for any preassigned r, the confidence level of the assertion that

$$1-\delta \leqq a/\theta \leqq 1+\delta,$$

where a is the minimax estimate of θ. This is important when r is small and asymptotic results are no longer valid.

Note. See also abstract No. **2**/674, **5.2**.

(A. P. Berens)

2/654

EPSTEIN, B. (Wayne State University and Stanford University) **4.4** (2.5)

Estimation of the parameters of the two parameter exponential distribution from censored samples—*In English*

Technometrics (1960) **2**, 403-406 (2 references)

The two-parameter exponential has the density function $f(x;\ \lambda,\ A) = \lambda \exp\{-\lambda(x-A)\}$, where $x \geqq A$, and A is, for example, a minimum "guaranteed" life span. In a context of life-testing, n items are placed on test and testing is stopped after r items have failed. The r failure times are $x_1 < x_2 < \ldots < x_r$: there is no replacement of these items. Let

$$T_k = \sum_{i=1}^{k} x_i + (n-k)\, x_k.$$

Then x_1 and $(T_r - nx_1)$ are independent and sufficient for estimating A and $\theta = 1/\lambda$.

Furthermore, $\hat{A} = x_1 - \theta/n$ and $\hat{\theta} = (T_r - nx_1)/(r-1)$ are unbiased and minimum variance for A and θ.

Confidence limits on θ are obtained from chi-square tables because $2(r-1)\,\hat{\theta}/\theta$ follows chi-square with $2(r-1)$ degrees of freedom. Confidence limits on A are determined from F tables since $n(x_1-A)/\theta$ and $2(r-1)\hat{\theta}/\theta$ are independent and chi-square with 2 and $2(r-1)$ degrees of freedom respectively; and thus this ratio is equal to $n(x_1-A)/\theta$ and follows F with 2 and $2(r-1)$ degrees of freedom. If $\hat{A}$ actually equals A, the ratio is unity; therefore the tabulated limits of F give limits on $n(x_1-A)/\theta$ and hence the α-limits on A are easily obtained.

The density function of $z = n(x_1-A)/\hat{\theta}(r-1)$ is $g(z) = (r-1)/(x+1)^r$, $z \geqq 0$. These can also be used to find confidence limits on A.

The paper closes with a numerical example.

(E. H. Lehman, Jr.)

2/655

FEND, A. V. (Stanford Research Institute, California) **4.1** (4.3)

On the attainment of Cramér-Rao and Bhattacharyya bounds for the variance of an estimate—*In English*

Ann. Math. Statist. (1959) **30**, 381-388 (5 references)

The results of this paper may be applicable in problems of minimum variance unbiased estimation. Suppose it is desired to estimate a real parameter θ, or some function thereof, from some data X having a (joint) density with a single unknown parameter θ. The Cramér-Rao bound or the Bhattacharyya bounds [see Cramér, *Mathematical Methods of Statistics* (1946) Princeton and Bhattacharyya, *Sankhyā* (1946) **8**, 1-32] on the variance of an estimator are sometimes useful for verifying whether or not an unbiased estimator has uniform minimum variance.

The author first proves, under common regularity conditions for these bounds, that the Cramér-Rao bound is achieved if and only if the logarithm of the density is of the form $t(x)\,g(\theta) + g_0(\theta) + h(x)$ where t is the estimator and g is not constant. He extends this result to prove that, if the density is of the Darmois-Koopman exponential family, and if the variance of the estimator attains the kth order Bhattacharyya bound but not the $(k-1)$th, then the logarithm of the density is of the form given above and the estimator is a polynomial in t of degree k. In fact, the variance of any kth degree polynomial in t will achieve the kth bound so that any such polynomial is a uniform minimum variance unbiased estimator of its expectation.

He points out that an unbiased estimator may achieve the kth order bound without being a maximum likelihood estimator, contrary to the case of the first order bound; that is the Cramér-Rao bound. He treats the example of a single observation x for which $x\theta^{-c}$ $(c>0)$ has a negative exponential distribution on the positive real numbers. If $1/c$ is an integer k, the kth bound is achieved by $x_k/k!$, the only unbiased estimator of θ, whereas the maximum likelihood estimator is x_k; if c is integral (>1), no polynomial is an unbiased estimator of θ, so that no unbiased estimator of θ can achieve any Bhattacharyya bound. In fact, by Laplace transform theory, $x^{1/c}/\Gamma(1+1/c)$ is the unique, and hence uniform minimum variance, unbiased estimator of θ.

(W. J. Hall)

2/656

GODAMBE, V. P. (Indian Statistical Institute, Calcutta) **4.9** (8.2)

An admissible estimate for any sampling design—*In English*

Sankhyā (1960) **22**, 285-289 (5 references)

For a finite population consisting of N individuals, on each of which a metric character X is defined, a fundamental problem is to estimate $T = \sum_{\lambda=1}^{N} X_\lambda$ on the basis of a sample. In an earlier paper [*J. R. Statist. Soc.* B (1955) **17**, 269-278] the author gave the most general definitions of the terms " sample " and " linear estimate ". With that terminology, he proves that the well-known unbiased estimator $\hat{T} = \Sigma\,(X_\lambda / P_\lambda)$, Horwitz & Thompson's estimator [*J. Amer. Statist. Ass.* (1952) **47**, 668-670], where λ runs through all distinct units in the observed sample and P_λ is the probability of the λth unit being included in a random sample is an admissible estimator, in the sense that there exists no other unbiased estimator which is uniformly superior to $\hat{T}$.

The author next proves that in the class of sampling designs with a fixed expected number of distinct units per sample, that is to say one which can be interpreted as a fixed expected cost of sampling, the estimator $\hat{T}$ has its maximum variance minimised, so that it possesses the minimax property.

(T. V. Hanumantha Rao)

2/657

GOODMAN, L. A. (University of Chicago) **4.3** (4.1)

A note on the estimation of variance—*In English*

Sankhyā (1960) **22**, 221-228 (10 references, 2 tables)

In this paper the author proves that if X is a random variable such that $\mathscr{E} \log (X \mid \theta)$ is a known constant A, then among all statistics of the form aX, the expression $e^{-A}X$ minimises $\mathscr{E}\,[\lambda(\theta)(\log aX - \log \theta)^2]$: it is also unbiased with respect to the loss function

$$[\log f(x) - \log \theta]^2.$$

A similar result was proved by the author in an earlier paper [*Ann. Math. Statist.* (1953) **22**, 114-117]. The loss function considered there was $\lambda(\theta)\,[f(x) - \theta]^2$. The conditions of the theorem are satisfied if the density function of X is of the form $\theta^{-1}\,g\,(x \mid \theta)$, $(\theta > 0)$.

The results mentioned above are applied to obtain best-unbiased estimates for the variance of a normal population. The loss functions considered are $(\sqrt{f} - \sigma)^2/\sigma^2$, $(\log \sqrt{f} - \log \sigma)^2$, $(f - \sigma^2)^2/\sigma^4$, $(f - \sigma^2)^2$ and $\Pr\,(\lvert f - \sigma^2 \rvert > k\sigma^2)$.

(S. John)

2/658

HARTER, H. L. (Aeronautical Research Laboratory) **4.4** (3.8)

The use of sample ranges and quasi-ranges in setting exact confidence bounds for the population standard deviation σ. I. The range of samples from a rectangular population —probability integral and percentage points; exact confidence bounds for σ—*In English*

ARL 31, Part I (1961) Aero. Res. Labs., Wright-Patterson Air Force Base.
xii+48 pp. (7 references, 3 tables)

The author discusses point estimates and interval estimates of the population standard deviation σ, based on the sample range and quasi-ranges. Harter [*Wright Air Development Center Technical Report* 58-200 (1958)] has tabulated the most efficient unbiased estimates of σ based on one quasi-range for samples of size $n = 2\,(1)\,100$ from normal, rectangular, and exponential populations and on linear combinations of two quasi-ranges for samples of size $n = 4\,(1)\,100$ from the normal population. Chu, Leone & Topp [*Ann. Math. Statist.* (1957) **28**, 530-531] have proposed a procedure, using sample quasi-ranges, of setting confidence bounds for the population standard deviation. Leone, Rutenberg & Topp [*Air Force Office of Scientific Research Report* 60-408 (1960)] have obtained approximate confidence bounds for the standard deviation of normal, exponential, and rectangular populations by first applying distribution-free methods and then imposing the distribution. Though the basic procedure is sound, the use of distribution-free methods yields approximate confidence intervals which are much longer than the exact confidence intervals which are given in the present report.

In the case of a rectangular population, the efficient point estimate and the most effective interval estimates are those based on the sample range, so it is not necessary to consider estimates based on sample quasi-ranges. The coefficients of the sample range w in the exact confidence bounds for the population standard deviation σ are found by taking the reciprocals of percentage points of the (standardised) range $W = w/\sigma$.

The following tables for the rectangular population are included:

(i) An eight-decimal-place table of the probability integral of the range for $W = 0{\cdot}01\ (0{\cdot}01)\ 3{\cdot}46\ (0{\cdot}001)\ 3{\cdot}464$ and sample sizes $n = 2\ (1)\ 20\ (2)\ 40\ (10)\ 100$;

continued

HARTER, H. L. (Aeronautical Research Laboratory) **4.4** (3.8) *continued*

The use of sample ranges and quasi-ranges in setting exact confidence bounds for the population standard deviation σ. I. The range of samples from a rectangular population —probability integral and percentage points; exact confidence bounds for σ—*In English*

ARL 31, Part I (1961), Aero. Res. Labs., Wright-Patterson Air Force Base.
xii+48 pp. (7 references, 3 tables)

(ii) a six-decimal-place table of the percentage points of the range corresponding to cumulative probabilities $P = 0{\cdot}0001$, $0{\cdot}0005$, $0{\cdot}001$, $0{\cdot}005$, $0{\cdot}01$, $0{\cdot}025$, $0{\cdot}05$, $0{\cdot}1\ (0{\cdot}1)\ 0{\cdot}9$, $0{\cdot}95$, $0{\cdot}975$, $0{\cdot}99$, $0{\cdot}995$, $0{\cdot}999$, $0{\cdot}9995$, $0{\cdot}9999$ for the same values of n;

(iii) a table, to seven significant figures or six decimal places, whichever is less accurate, of the coefficients of the sample range w in the exact lower confidence bounds for σ for the above values of P and n.

Work is now in progress on Part II of this report, which will give exact confidence bounds, based on sample quasi-ranges, for the standard deviation of a normal population.

(H. L. Harter)

MURTHY, M. N. & NANJAMMA, N. S. (Indian Statistical Institute, Calcutta) **4.1** (4.9)
Almost unbiased ratio estimates based on interpenetrating subsample estimates—*In English*
Sankhyā (1959) **21**, 381-392 (4 references)

In this paper, the efficiencies of different types of ratio estimates built up from n independent, interpenetrating subsamples have been compared with regard to their biases and mean-square errors. The main comparison is between the two types $R_1 = (y_1+\ldots+y_n)/(x_1+x_2+\ldots+x_n)$ and $R_n = (y_1/x_1+\ldots+y_n/x_n)/n$; the y_i's and x_i's being unbiased estimates of the mean values of y and x obtained from the ith subsample. It is shown that the bias of R_n is n times that of R_1 to the second degree of approximation. The mean-square error, M_n of R_n, is $M_1+n^{-2}(n-1)A+(1-1/n)B^2$ where B is a constant; A is another constant positive for bivariate normal distributions, and M_1 is the mean-square error of R_1 to the fourth degree of approximation.

Using these results an almost unbiased estimate $R_c = (nR_1-R_n)/(n-1)$ is constructed and conditions are derived for the variance of R_c to be less than the mean-square error of R_1 in terms of B, A and the correlation ρ between R_1 and R_n. Similarly, for some suitable values of ρ, the minimum value of n for which the gain in using R_c is greater than that in using R_1 is positive and the value of n for which the gain is a maximum are tabulated.

With the help of an empirical study conducted with $n = 2$, the validity of the theoretical results is demonstrated: generalising the results further, ratio estimates of the type R_m, obtained by grouping the n subsamples into m groups, are compared taking the approximation to any general $2k$th degree.

(G. Parthasarathy)

2/661

NANJAMMA, N. S., MURTHY, M. N. & SETHI, V. K. (Indian Statist. Inst., Calcutta) **4.9** (4.3)
Some sampling systems providing unbiased ratio estimators—*In English*
Sankhyā (1959) **21**, 299-314 (5 references)

This paper deals with sampling designs that provide unbiased ratio estimators in the case of sampling from a finite population. For example, selecting the first unit with probability proportional to x, which is known for every unit, and the rest of the $(n-1)$ units from the remaining $(N-1)$ units with equal probability without replacement leads to the unbiased ratio estimator $\bar{y}/\bar{x}$. An estimate of its variance is derived and it is shown that this unbiased ratio estimator is more efficient than the corresponding biased one which can be obtained by simple random sampling of n units from N units, if the correlation between $\bar{y}^2/\bar{x}$ and $\bar{x}$ is not negative. In the case of a stratified one-stage design, $(\Sigma N_i\bar{y}_i)/(\Sigma N_i\bar{x}_i)$ becomes unbiased if one unit out of all the units from all the strata is selected with probability proportional to x; the remaining (n_i-1) units from that stratum which contains the selected unit and all the n_j units from the jth stratum $(j \neq i)$ are selected with equal probability without replacement from the respective strata.

As a generalisation, a necessary and sufficient condition has been derived for estimating a general class of parameters F, expressible as $\sum_{\alpha \in A} f(\alpha)$ where α is a subset of the population; A is the class of sets α, and $f(\alpha)$ a set function. F is estimable if and only if each sample w contains at least one set α and each α is contained in at least one w. Then

$$\hat{F} = [\sum_{\alpha \in w} f(\alpha)\,\phi(w, \alpha) \div \Pr(w)]$$

where $\phi(w, \alpha)$ is such that $\sum_{w \supset \alpha} \phi(w, \alpha) = 1$ for all α, is an unbiased estimator F. An estimate of the variance of F is also given. By taking $\phi(w, \alpha) = \Pr(w \mid \alpha)$, the conditional probability of getting the sample w given that the test α has been selected first, the usual estimators are obtained in many cases. In order to obtain an unbiased estimate of F/G, where G is another parameter, it is sufficient first to select the set α with probability proportional to $g(\alpha)$ and the rest of the units with any probability scheme. In this case, the unbiased ratio estimator is given by

$$\hat{R} = \sum_{\alpha \in w} f(\alpha) \Pr(w \mid \alpha) / \sum_{\alpha \in w} g(\alpha) \Pr(w \mid \alpha).$$

The estimate of its variance is also derived. Applications of these generalised results are made to estimate population mean, variance and the regression coefficients: the procedure is extended to two-phase sampling.

(G. Parthasarathy)

2/662

NARAYANA, T. V., CHORNEYKO, I. & SATHE, Y. S. (Univ. Alberta, Canada) **4.7** (4.3)
Sufficient partitions for a class of coin tossing problems—*In English*
Biom. Zeit. (1960) **2**, 269-275 (9 references, 1 figure)

The authors consider estimation problems in a Markoff-chain of alternatives. Let the probability of a coin falling head be p_1 (be p_2), if in the previous trial the outcome was tails (was heads). At the first trial the probability of a head is p_1. Two different stop-rules are considered: first, the coin is tossed until the total number of heads exceeds the total number of tails by r, and second the coin is tossed n times.

(i) If the first trial results in tails, let n be the total number of tails and let m be the number of changes from tails to heads or, if the first trial results in a head, let n again be the total number of tails, but $m-1$ be the number of changes from tails to heads. It is shown that (n, m) is a sufficient statistic and the minimal sufficient partition in the sample-space is constructed. The statistic (n, m) is complete, so there exists a unique unbiased estimator for p_1, which has minimum variance. This estimator and also the estimator for p_2 are given.

(ii) Let r be the number of heads in n tosses. We define $R = +r$, if the last trial resulted in a head and $R = -r$, if the last trial resulted in tails. Let k be the number of changes from tails to heads. Then (R, k) is a sufficient statistic and it is shown, that it is also complete. Finally the authors give the minimum unbiased estimator for p_1.

(W. Vogel)

2/663

PLACKETT, R. L. (University of Liverpool) **4.3** (3.8)
The analysis of life test data—*In English*
Technometrics (1959) **1**, 9-19 (15 references, 5 tables, 3 figures)

A common problem in life testing is the estimation of the probability that the life, L, exceeds a specified value L_0. Estimation procedures are given, without derivation, which are in current use. An introduction is given to the ideas behind order statistics, with illustrative examples from the normal and logistic distributions. A graphical method is then discussed for estimating the parameters of these distributions from singly censored samples.

A second quick estimation procedure for estimating the mean and standard deviation from a small normal sample involves equating the lowest and highest censored values to their respective expectations and solving for the unknown parameters; this technique has good efficiency for very small sample size.

Exact techniques, developed by Gupta [*Biometrika* (1952) **39**, 260-273], and expanded by Sarhan and Greenburg [*Ann. Math. Statist.* (1956) **27**, 427-451 and (1958) **29**, 75-105], give the appropriate coefficients for best linear unbiased estimators of μ and σ, based on observations in a small ($n \leqq 15$) censored sample from a normal population. An alternative technique, applicable for $n \leqq 50$ which is the present limit for available tables of normal order statistics, is presented which does not depend upon the tables of Gupta and other writers; and which employs only tables of the normal distribution and tables of the expected values of normal order statistics. The latter technique is also illustrated as applied to the logistic distribution.

Finally, an approximate technique to handle large ($n > 50$) samples from a normal population is discussed, in which the observations are grouped into frequency classes. The normal equivalent deviations for each midpoint is taken to represent the average expectation of the normal order statistic, after which the technique described for small ($n \leqq 50$) normal samples is applied.

(H. E. McKean)

2/664

STEIN, C. (Stanford University) **4.1** (4.3)
The admissibility of Pitman's estimator of a single location parameter—*In English*
Ann. Math. Statist. (1959) **30**, 970-979 (10 references)

Pitman [*Biometrika* (1939) **30**, 391-421] has presented an estimator for a location parameter, based on a random sample with a single unknown parameter, which has uniformly minimum mean-square error among all translation invariant estimators. This estimator has later been proved to be minimax [Girshick & L. J. Savage, *Proc. Second Berk. Symp.* (1951) 53-73], without any invariance restriction; and Karlin [*Ann. Math. Statist.* (1958) **29**, 406-436], under fairly strong conditions on the form of the density function, has proved it to be admissible in terms of mean-square error. This paper presents an alternative, and less restrictive proof of this admissibility.

The method is a generalisation of one due to Blyth and requires essentially that Pitman's estimator have a third moment. Thus, for example, for samples of size at least seven, the sample median is an admissible estimator of the median of a one-parameter Cauchy distribution.

(W. J. Hall)

2/665

STÖRMER, H. (Siemens & Halske, Munich) **4.4** (2.5)
On a test for the parameters of a life distribution—*In German*
Metrika (1961) **4**, 63-77 (1 reference, 1 table, 1 figure)

The author considers the distribution function

$$f(t) = 1 - \exp\{-(t/T)^{\alpha}\}$$

which occurs with life-studies. On account of the difficulty of obtaining estimates for α and T in the usual way, the following procedure is introduced. A number of items (N) are observed; as soon as an item fails it is substituted by another item of equal age, so that the number of items under observation is constant all the time. Let $t_1, \ldots, t_n$ be the ages of the first n items which fail. By considering the joint distribution of $(t_1, \ldots, t_n)$, it is seen that with $u_1 = \sum_{i=1}^{n-1} \ln t_n/t_i$ and $u_2 = t_n$, (u_1, u_2) is a sufficient statistic. The maximum-likelihood method leads to the estimates $\hat{\alpha} = n/u_1$ and $\hat{T} = (N/n)^{u_1/n}\mu_2$.

Finally, the variables $\alpha/\hat{\alpha}$ and $(T/\hat{T})^{\alpha}$ are introduced which have a distribution function independent of α and T. The marginal-distribution of the first variable is closely related to the χ^2-distribution and $2n\alpha/\hat{\alpha}$ is distributed as $\chi^2_{2(n-1)}$. The marginal-distribution of the second variable seems to be in no immediate relation to one of the usual distribution functions.

(J. Pfanzagl)

2/666

THEIL, H., & SCHWEITZER, A. (Netherlands School of Economics, Rotterdam) **4.3** (6.0)
The best quadratic estimator of the residual variance in regression analysis—*In English*
Statist. Neerlandica (1961) **15**, 19-23 (2 references)

This article discusses the estimation of the variance σ^2 of the disturbances in regression analysis under classical assumptions: the " descriptive " variables take non-stochastic values and the disturbances are normal, independent and homoscedastic.

It is known that the maximum likelihood estimate is found by determining the sum of the squares of the disturbances estimated by least-squares and dividing this sum by T, the number of observations; also, that the best quadratic unbiased estimator of σ^2 is obtained by dividing the sum of squares of the disturbances estimated by least-squares by $T-\lambda$, where λ represents the number of coefficients to be estimated; i.e. the number of characteristic variables ($+i$ if a constant term is present). It is shown that the best quadratic estimator is obtained by dividing the sum of squares of the disturbances estimated by least-squares by $T-\lambda+2$.

(H. Theil)

2/667

TRYBUŁA, S. (Wrocław) **4.0** (0.7)
On some loss functions—*In English*
Coll. Mat. (1960) 7, 297-305

Two statistical games are considered in which minimax estimators are found to be unbiased.

The first deals with the multinomial distribution and may be defined in terms of an urn scheme as follows: $m = (m_1, \ldots, m_k)$ is the result of drawing with replacement n balls from an urn containing balls of k different colours, m_i representing the number of balls of the ith colour in the sample. A vector function $f = (f_1, \ldots, f_k)$, where $f_i = f_i\,(m)$, is an estimator of the unknown composition $p = (p_1, \ldots, p_k)$ of the urn, and constitutes a strategy of the statistician. The prior distribution of p is the strategy of his opponent. The author shows that if the loss function is

$$\sum_{i=1}^{k} (f_i - p_i)^2/p_i,$$

then the minimax estimator is $f = (m_1/n, \ldots, m_k/n)$, the game is determined and uniform distribution over the $(k-1)$-dimensional range of p is the least favourable prior distribution.

The author also derives a similar result for the hypergeometric distribution. Here it is assumed that an urn contains a fixed number N of balls of k different colours, and the statistician wants to estimate how many balls of particular colours there are in the urn, knowing the composition of a sample of size n drawn without replacement. The author finds a loss function for which the unbiased estimator $(Nm_1/n, \ldots, Nm_k/n)$ turns out to be the minimax estimator of the content $M = (M_1, \ldots, M_k)$ of the urn, the uniform distribution of M over its range to be the least favourable prior distribution, and the statistical game in question to be determined.

(S. Zubrzycki)

2/668

VRIES, M. de. (A. C. Nielson, Amsterdam) **4.0** (6.0)

Practical estimators for sampling on successive occasions—*In Dutch*

Statist. Neerlandica (1961) **15**, 123-133 (3 references)

The sample size in a sample investigation covering a length of time may vary from period to period. A number of estimation methods are suggested by which information from previous observations is utilised as much as possible. These methods are compared with the one for estimating the regression which is theoretically optimal. When observing a large number of characteristics the regression estimate can no longer be used because of the large number of calculations involved. The estimators suggested are less precise but do not cause any technical calculation difficulties.

(M. de Vries)

2/669

WALSH, J. E. (System Development Corporation, Santa Monica, Cal.) **4.6** (5.6)

Nonparametric linear estimation of common median of symmetrical populations from symmetrically censored observations—*In English*

Sankhyā (1960) **22**, 295-300 (5 references)

The author is concerned with the problem of estimating the median of a symmetrical population using the observations in a symmetrically censored sample. The underlying data consist of *n* observations whose ordered values are denoted by $x_{(1)} \leqq x_{(2)} \leqq \cdots \leqq x_{(n)}$. In the available set of observations only $x_{(r)}$, $x_{(r+1)}$, ..., $x_{(n+1-r)}$ are known.

The author considers linear estimates of the form

$$\sum_{r}^{n+1-r} a(i)\, x_{(i)}$$

where $a(i) = a(n+1-i)$ for all i,

$$\sum_{r}^{n+1-r} a(i) = 1,$$

and proposes some heuristic working rules for an efficient determination of the coefficients $a(i)$ based on the shape of the distribution. These rules can be applied to the discrete case on the basis of the shapes of histograms for discrete frequency functions and provide valid estimates for the median even when the symmetrical population sampled is not of the specified shape; or when the observations come from different symmetrical populations with the common median ϕ. These rules were suggested to the author from Sarhan's papers on best linear estimation of the mean and standard deviation in populations of specified parametric form using available order statistics [*Ann. Math. Statist.* (1955) **26**, 505-511 and 576-592].

For the special case of no censoring ($r = 1$) the author suggests the following nonparametric test for the hypothesis $\phi = \phi_0$. When the hypothesis is true and only the absolute differences of the sample observations from ϕ_0 are given, all possible samples which satisfy these conditions have conditionally the same probability of occurrence. If $\hat{\phi}$ be an estimator for the median determined by the working rules stated, it is proposed that $\hat{\phi} - \phi_0$ could be used as a test criterion for the hypothesis $\phi = \phi_0$ and that its level of significance be determined from the conditional distribution indicated above. A large-sample approximation to this procedure is obtained.

Finally a method is stated for obtaining confidence intervals for the unknown median on the basis of these tests.

(S. K. Mitra)

2/670

WEISSBERG, A. & BEATTY, G. H. (U.S. Food & Drug Admin., Washington, & Batelle Memorial Inst., Columbus, Ohio) **4.5** (11.1)

Tables of tolerance-limit factors for normal distributions—*In English*

Technometrics (1960) **2**, 483-500 (15 references, 2 tables, 1 figure)

The aim of this paper is to present tables permitting an experimenter to predict that at least a given percentage P, of future measurements from a normal universe will with a specified probability γ, lie within a computed interval $\bar{X} \pm Ks$, where $\bar{X}$ and s are conventionally defined and $K(N, P, f, \gamma) = ru$.

Define $r(N, P)$ in this way:

$\Psi(1/\sqrt{N}+r)-\Psi(1/\sqrt{N}-r) = P$, where N is the sample size and $\Psi(x)$ is the probability that a standardised normal variate does not exceed x.

Define $u(f, \gamma)$ in this way: $u^2 = f/\chi^2_{f,\gamma}$, $u \geqq 0$, where f, not necessarily equal to $N-1$ is the degrees of freedom for s and $\chi^2_{f,\gamma}$ is the abscissa on the χ^2 density function with f degrees of freedom such that γ of the distribution lies to the right.

Newton's method, $r_{i+1} = r_i - [f(r_i)/f'(r_i)]$, was used in order to compute r. Here

$$f(r) = \Psi(1/\sqrt{N}+r)-\Psi(1/\sqrt{N}-r)-P = 0.$$

Table r gives r for $N = 1$ (1) 100 (5) 200 (10) 300 (20) 500 (100) 1000, 2000, 3000, 5000, 10,000, ∞, and $P = 0{\cdot}50, 0{\cdot}75, 0{\cdot}90, 0{\cdot}95, 0{\cdot}99, 0{\cdot}999$.

Calculations from Thompson's, Hald and Sinkbaek's, and Pearson's tables were required to compute u. Table u gives u for $f = 1$ (1) 150 (2) 250 (5) 500 (10) 800 (20) 1000 (1000) 10,000, ∞, and $\gamma = 0{\cdot}90, 0{\cdot}95, 0{\cdot}99$.

Calculations of r and u were performed on the IBM 650 at Battelle with an accuracy of at worst one unit in the fourth decimal place, except for $u\,(f<20, \gamma = 0{\cdot}999)$ which is uncertain.

Now since $K = ru$, given N, P, F, and γ we may uniquely determine K: were we to list tables for $K(N, P, f, \gamma)$ directly, about 1660 times as many entries would be required. Bowker has given a table similar to this but requiring $f = N-1$.

The authors point out an imperfection in their tables. The interval $\bar{X}+rus$ really corresponds to N, P, f, and γ'. The exact value $|\gamma'-\gamma|$ is unknown, but Wald & Wolfowitz have given some typical limits on γ. The authors include the estimated number of significant decimal places in the product ru for selected combinations of N, P, f, and γ.

2/671

(E. H. Lehman, Jr.)

WORMLEIGHTON, R. (University of Toronto, Ontario) **4.7** (4.4)

A useful generalisation of the Stein two-sample procedure—*In English*

Ann. Math. Statist. (1960) **31**, 217-221 (1 reference, 1 figure)

A general two-sample procedure is proposed for estimating the mean, μ, of a normal population with unknown variance σ^2. Special cases discussed include:

(i) fixed sample size,
(ii) Stein's two-sample method, in which the precision of the estimate is fixed ["A two-sample test for a linear hypothesis whose power is independent of the variance", *Ann. Math. Statist.* (1945) **16**, 243-258],
(iii) a modified Stein procedure with an upper bound on the total number of observations,
(iv) a method in which the decisions on sample size and precision are postponed until after the first sample is obtained, and
(v) a procedure which minimises the total cost, i.e. the cost of the observations plus the cost resulting from a lack of precision of the estimate.

The procedure is as follows: let s^2 be the sample variance of a first sample of n_0 observations. Let n be the total number of observations, let $2L$ be the length of a $(1-\alpha)$ confidence interval for μ, and let $t^{(\alpha)}$ be the $(1-\alpha)$ point of a t-variate with (n_0-1) degrees of freedom. Consider the family of curves, indexed by s,

$$n_s(L) = (st^{(\alpha)}/L)^2.$$

(*a*) On each curve choose in advance, a single point: this set of points constitutes a "cut" across the family.

(*b*) Take the first sample and calculate s: this determines a particular curve of the family and, on the cut, a unique point, (n^*, L^*).

(*c*) Take $[n^*-n_0]+1$ further observations, where $[q]$ denotes the largest integer strictly less than q.

(*d*) Calculate $\bar{x}$, the mean of all the observations: then $\bar{x} \pm L^*$ is a $(1-\alpha)$ confidence interval for μ.

The proof is that given by Stein. Appropriate definitions of the cut lead to the special cases discussed.

(R. Wormleighton)

2/672

ELTEREN, P. van (Catholic University, Nijmegen) **5.7** (1.8)

Two restricted sequential sign tests—*In English*

Statist. Neerlandica (1961) **15**, 91-96 (3 references, 3 figures)

During the last few years Armitage, " Restricted Sequential Procedures " [*Biometrika* (1957) **44**, 9-26] has stimulated the use of sequential methods which limit the number of observations required to arrive at a decision. There also exist restricted versions of the classic sequential schemes of Wald. Here a sequential sign test of Armitage with three decision possibilities is compared with the corresponding Wald test which is further discussed by de Boer. The Wald scheme is a sequential test with three possible decisions for testing an unknown probability. It is shown that the efficiency of the Wald test is dependent more on the order in which the results are obtained than the Armitage test so that, in many cases, the latter will be preferable.

(P. van Elteren)

2/673

EPSTEIN, B. (Wayne State University and Stanford University) **5.2** (2.5)

Tests for the validity of the assumption that the underlying distribution of life is exponential. I.—*In English*

Technometrics (1960), **2**, 83-101 (31 references)

This article deals with a number of graphical and analytical procedures for testing, on the basis of life test data, the assumption that the underlying distribution of life is exponential. Part I contains descriptions of the mathematical and graphical procedures, while Part II [*Technometrics* (1960) **2**, 167-183] contains many worked examples illustrating the methods of Part I.

Included are 12 tests, labelled by Epstein as follows:

(1) a graphical procedure,
(2) the χ^2 test for goodness of fit,
(3) a criterion based on the conditional distribution of total lives,
(4) a test for abnormally early failures,
(5) a test for an abnormally long first failure,
(6) a test for whether the mean-life, or rate of failure in the first half of a life test differs significantly from the mean-life or rate of failure in the second half of a life test,
(7) a test for whether or not the mean-life fluctuates during the life test,
(8) a special case of (7) when the group size is unity,
(9) a test based on the maximum F distribution,
(10) tests for abnormally long periods in which there is no failure,
(11) a graphical procedure based on the Kolmogoroff-Smirnoff test, and
(12) a test based on the conditional rate of failure.

The description of these tests is followed by general comments, a short summary, an appendix in which some of the more mathematical arguments are found, and finally by a bibliography.

Some of the tests are restricted in various ways which are clearly noted by the author. Several remarks are given following each test indicating variations on that particular test.

Note. See also abstract No. **2**/654, **4.4.**

(R. L. Conn)

2/674

KLOTZ, J. (University of California, Berkeley) **5.6** (5.8)

Null distribution of the Hodges bivariate sign test—*In English*

Ann. Math. Statist. (1959) **30**, 1029-1033 (1 reference, 1 table)

For measurement vectors (x_i, y_i) (x_i', y_i'), $i = 1, \ldots, n$, from two bivariate distributions, Hodges [*Ann. Math. Statist.* (1955) **26**, 523-527] proposed a sign test for the hypothesis that the bivariate distributions are the same and that the vectors are independent. He considered the vectors of differences $(x_i'-x_i, y_i'-y_i)$ and the maximum value $(= M)$ of the number of projected vectors with a given sense over all possible directions. The hypothesis is then to be rejected if M is too large. The particular combinatorial problem of determining the number of sign patterns was then resolved by identifying it with the classical gambler's ruin problem. Hodges gave expressions for the number of ways $m_h(n)$ of proceeding from the point (0, 0) to $(k, n-k)$ and then hitting the line: $y = x+h$ only at the nth step. Values of $\Pr\{K \leqq k\}$ for the number of patterns with at most k negative signs uppermost were computed for $n = 1\,(1)\,30$ in case $k < n/3$.

The present paper extends the values of $\Pr\{K \leqq k\}$ for $n = 31\,(1)\,50$ and provides values of P for k such that P just exceeds 0·10. Expressions for $\Pr\{K = k\}$ are also derived in the more general case:

$$(2j+1)\,h \leqq n < (2j+3)\,h.$$

(B. M. Bennett)

2/675

PFANZAGL, J. (Universität Köln) **5.0** (5.1)

On the existence of uniformly most powerful tests—*In German*

Metrika (1960) **3**, 169-176 (4 references)

Let $\{P_\theta, \theta \in \Omega\}$ be a class of probability measures on a measurable space (X, A). Assume that each P_θ is absolutely continuous with respect to a σ-finite measure μ. Denote the Radon-Nikodym derivatives of P_θ with respect to μ by $p_\theta(x)$. One of the probability measures of the class $\{P_\theta\}$ is marked as the null hypothesis P_0, the remaining P_θ form the alternative hypotheses. The quotient of $p_\theta(x)$ divided by $p_0(x)$ is called the likelihood ratio $f_\theta(x)$. It is well known that the monotony of these likelihood ratios is sufficient for the existence of a uniformly most powerful test for P_0 against all alternatives P_θ for every level of significance α.

In the present paper it is proved that this condition of monotony is also necessary.

For every level of significance α $(0 \leqq \alpha \leqq 1)$ there exists a uniformly most powerful test statistic $\phi(x, \alpha)$, defined for almost every $x \in X$, for the null hypothesis P_0 against all alternatives P_θ. Then there exists a real valued, measurable function $T(x)$, depending only on P_0 and ϕ, and a class of real monotonic functions H_θ with the property that the likelihood ratios $f_\theta(x)$ can be represented in the form $f_\theta(x) = H_\theta T(x)$ for P_0, almost every $x \in X$ and all $\theta \in \Omega$. $T(x)$ is a sufficient statistic. The relation to a theorem of Neyman & Pearson [*Statist. Res. Mem.* (1936) **1**, 113] is mentioned.

(W. Uhlmann)

Remark of the Regional Editor:

There is one weak point in the proof concerning the independence of $T(x)$ from the alternative; a supplementary note will appear in *Metrika* (1961) **4**, issue No. 2.

(J. Pfanzagl)

2/676

ROY, J. (Indian Statistical Institute, Calcutta) **5.8** (3.9)
Non-null distribution of the likelihood ratio in analysis of dispersion—*In English*
Sankhyā (1960) **22**, 289-294 (8 references, 3 tables)

It is well known that every problem in analysis of dispersion, that is the generalised analysis of variance, can be put in the following form: a random matrix Y of order $(p \times m)$ of uncorrelated p-dimensional column vectors is given; and it is known that all the column vectors have the same dispersion matrix. An unbiased estimate V of the dispersion matrix is available. In order to test whether $\mathscr{E}(Y) = 0$, we assume that the joint distribution of the elements of Y is multivariate normal and that the estimate of the common dispersion matrix of the column vectors is distributed independently of Y according to the Wishart's law with n degrees of freedom. The appropriate likelihood ratio criterion is $\Lambda = |W| / |W+YY'|$ where $W = nV$. The null distribution of this criterion has been investigated extensively, see for instance, [C. R. Rao, *Advanced Statistical Methods in Biometric Research* (1952) New York: Wiley].

The present paper obtains an asymptotic expansion of the non-null distribution of Λ in the case where $\mathscr{E}(Y)$ is of rank one. The expansion is in terms of the distribution function of the chi-square variable. The author observes that the sum of the first few terms in the expansion gives satisfactory results provided the deviations in the parameters are small and that n is large. A table of the power function of the analysis of dispersion test at the 5 per cent. level of significance is given for $p = 1\,(1)\,4$; $m = 2, 3$ and $n = 200$.

(S. John)

2/677

UZAWA, H. (Stanford University, California) **5.6** (5.3)
Locally most powerful rank tests for two-sample problems—*In English*
Ann. Math. Statist. (1960) **31**, 685-702 (7 references)

Let $X_1, \ldots, X_{n_1}, X_{n_1+1}, \ldots, X_n$ be independent random variables where, for $1 \leqq i \leqq n_1$, X_i is distributed according to a distribution function F, and for $n_1+1 \leqq i \leqq n$, X_i is distributed according to a distribution function G. Let Z_i be zero or one according to whether the ith smallest member of the combined sample is distributed according to F or G. The author discusses the problem of characterising tests of the hypothesis $F = G$ which are based on $Z_1, \ldots, Z_n$ and which are locally most powerful in the sense that there is a one-parameter family of alternatives $\{(F_\theta, G_\theta)\colon 0 \leqq \theta \leqq \theta_1\}$ such that $F_0 = G_0$ and the test is uniformly most powerful against $\{(F_\theta, G_\theta)\colon 0 \leqq \theta \leqq \theta_0\}$ for some $0 < \theta_0 \leqq \theta_1$.

The form of any locally most powerful rank test is "reject the hypothesis if $\sum_{i=1}^{n} a_i Z_i > c$, and accept the hypothesis if the opposite inequality holds". The constants a_i are given explicitly.

Criteria are given for four problems to enable one to decide if a test of the above form is locally most powerful against some one-parameter family of alternatives. In order to test $F = G$ against the alternative $F \neq G$, then any test of the above form is locally most powerful. Every such test is locally most powerful against a one-parameter family having the property that $F_\theta G_\theta^{-1}(x) = (1-\theta)\,x + \theta H(x)$, where H is a distribution function on the closed unit interval.

To test $F = G$ against the alternative $F > G$, a rank test is locally most powerful if and only if it is of the stated form and, in addition, $(c_0, c_1, \ldots, c_{n-2})$ has all non-negative Hankel determinants, where

$$c_j = \left[1 \Big/ \binom{n-2}{j}\right] \sum_{s=j}^{n-2} \binom{s+1}{j+1} (a_{s+2} - \bar{a}).$$

If the problem is to test $F = G$, symmetric, against the alternative $F \neq G$, symmetric with the same median, a rank test is locally most powerful if and only if it is of the stated form and, in addition

$$\sum_{s=j+1}^{n} \binom{s-1}{j} (a_{n+1-s} - a_s) = 0$$

for $j = 1, 2, \ldots, n$. To test $F = G$ against the alternative $\int F\,dG > \tfrac{1}{2}$, a rank test is locally most powerful if and only if it has the form mentioned with

$$\sum_{s=1}^{[(n+1)/2]} ((n+1)/2 - s)(a_{n+1-s} - a_s) \geqq 0.$$

(D. R. Truax)

2/678

WIJSMAN, R. A. (University of Illinois, Urbana) **5.7** (5.1)

A montonicity property of the sequential probability ratio test—*In English*

Ann. Math. Statist. (1960) **31**, 677-684 (6 references)

Based on the inequality

$$aP_1\,(a<Y_N<b) \leqq P_2\,(a<Y_N<b) \leqq bP_1\,(a<Y_N<b)$$

where Y_N is the sequential probability ratio at the *n*th observation, the indices 1, 2, refer to the two hypotheses under consideration, and a and b are any real numbers, “ it is shown that, if, in a sequential probability ratio test, the upper stopping bound is increased and the lower stopping bound decreased, and if the new test is not equivalent to the old one, then at least one of the error probabilities is decreased ”. This result implies a previous result of Weiss [*Ann. Math. Statist.* (1956) **27**, 1178-1181] concerning the monotonicity of the error probabilities as a function of the stopping bounds and generalises the result of T. W. Anderson & Friedman [*Contributions to Probability and Statistics. Essays in Honor of Harold Hotelling* (1960) Stanford: University Press] concerning the optimum properties of the sequential probability ratio test.

In the words of the author, “ The monotonicity property is a consequence of the following stronger result. Let the old and new tests be given by the stopping bounds (B', A') and (B, A), respectively, with $B<B'<A'<A$; let (α_1', α_2') and (α_1, α_2) be the error probabilities and $\Delta\alpha_i = \alpha_i - \alpha_i'$ the changes in the error probabilities; then the vector $(\Delta\alpha_1, \Delta\alpha_2)$ is restricted to a cone consisting of the third quadrant, plus the part of the second quadrant where $-\Delta\alpha_2/\Delta\alpha_1 < B$, plus the part of the fourth quadrant where $-\Delta\alpha_2/\Delta\alpha_1 > A$. Another consequence of this result is that (α_1, α_2) cannot lie in the closed triangle with vertices (α_1', α_2'), (0, 1) and (1, 0) ”.

“ Finally, the following monotonicity property follows: if the lower stopping bound is fixed and the upper stopping bound increased, then $\alpha_1/(1-\alpha_2)$ decreases monotonically. The same holds for $\alpha_2/(1-\alpha_1)$ if the upper stopping bound is held fixed and the lower stopping bound decreased.”

(C. A. Bennett)

2/679

ZAJTA, A. (Agricultural University, Budapest) **5.3** (5.6)

On the Lehmann test—*In Hungarian*

Publ. Math. Inst. Hung. Acad. Sci. (1960) **5**, 447-495 (3 references, 1 table)

Lehmann, see “ Consistency and unbiasedness of certain non-parametric tests [*Ann. Math. Statist.* (1951) **22**, 165-180] and Rényi [*Magy. Tud. Akad. Alk. Mat. Int. Közl.* (1953) **2**, 243-265], has modified the Wilcoxon test into a form which is consistent against all alternative hypotheses. Referent, “ On two modifications of the Wilcoxon test ” [*Publ. Math. Inst. Hung. Acad. Sci.* (1959) **4**, 313-319], has shown the identity of these modifications to be a simple linear relationship between their statistics.

In this paper a new proof of this connection is given and a new expression of the statistic is established. This new expression is, in certain respects, simpler than the earlier ones as it is only a linear function of the rank numbers.

The variance of the statistic is computed in the general case of alternatives. As a special case the variance in the case of null hypothesis is also obtained.

(E. Csáki)

2/680

BASMANN, R. L. (Hanford Laboratories Operation, G.E.C., Richland, Washington) **6.6** (6.1)
An expository note on estimation of simultaneous structural equations—*In English*
Biometrics (1960) **16**, 464-480 (27 references, 2 tables)

The purpose of this paper is to make known to biologists and biometricians several methods of generalised linear estimation of structural parameters not yet widely known outside econometrics; in particular, the least-variance difference method. The technique of least-variance difference estimation is formulated exactly as the generalisation by Aitken [*Proc. Roy. Soc. Edinb.* (1934-35) **55**, 42-48] of Gauss's method of treating independent observations of unequal precision to the case of interdependent observations.

The author begins with a discussion of some logical features of a simple causal path model of the type presented by Turner & Stevens [*Biometrics* (1959) **15**, 236-258: abstracted in this journal No. **1**/258, **6.6**]. The model is expressed functionally in two equivalent ways, reduced form and structural form. The hypothesis that the structural equations are linear in the unknown parameters and some tests that might be made are considered. The author then discusses three of the several ways one may obtain the least-squares estimates of the structural parameters which are needed, these methods differing only in the degree to which estimates of the coefficients involved are constrained to satisfy the hypothetical restrictions deduced from the structural model. It is emphasised that the estimation processes described are conceptual only; the computational processes, however, are equivalent to the conceptual substitution procedure used. The three methods considered in this paper lead to the full-information estimates, the least-variance ratio estimates using the limited-information single equation method, and what the author calls the least-variance difference estimates.

In continuing, the author describes the heuristic considerations that led to his initial formulation of the least-variance difference method, these considerations being inspired by the classical Gauss-Fisher least-squares problem. This formulation involves nothing more than a slight generalisation of the problem that motivated the Gauss-Legendre formulation of the method of least-squares itself and could serve as an introduction to the broader subject of this article.

A thorough discussion of the statistical properties of the estimates yielded by the three methods follows. The paper ends with a numerical example of the least-variance difference method using artificial data and a test of the hypothesised relation by two structural equations. (S. R. Knox)

2/681

BASMANN, R. L. (Hanford Labs. Operation, G.E.C., Richland, Washington) **6.6** (4.8)
On the asymptotic distribution of generalised linear estimators—*In English*
Econometrica (1960) **28**, 97-107 (22 references)

This paper gives a sufficiently general proof of the asymptotic normality of the distribution of the generalised classical linear estimators of a single equation in a system of structural equations: see also T. W. Anderson & Rubin [*Ann. Math. Statist.* (1949) **20**, 46-63 and (1950) **21**, 570-582]. The author here makes the assumption that the random structural disturbances and the reduced-form disturbances are not normally distributed. The generalised classical linear estimators: for example the minimum variance-ratio estimators are not maximum-likelihood estimators. The effect of this is to require proof of the property of asymptotic normality under non-normal assumptions.

The second section of the paper states the stochastic and structural assumptions: use is made of matrix methods here. The next section deals with linear decompositions of the general classical linear estimators. After dealing with the asymptotic normality of two of the estimators in section 4 the final section considers a large sample test of identifiability.

(W. R. Buckland)

2/682

BATSCHELET, E. (Basle University) **6.9** (4.3)
On a contingency-table with missing data—*In German*
Biom. Zeit. (1960) **2**, 236-243 (3 references, 3 figures)

Let f_{ji} be the observed frequencies of the contingency-table, where the first index j denotes the row and the second index i the column of the table. The total number of columns is denoted by k and the number of rows by l. The frequencies of a column are supposed to be proportional up to random deviations; that is, say homogeneity exists between the rows, and may represent the marginal probability p_j. Further, it is supposed that in the first row only m frequencies, in the second only $m+n$, etc. are available, where $m<m+n<\ldots<k$. The problem is to evaluate the maximum likelihood estimates of the marginal probabilities p_j.

The author calculates these estimates and points out that they are proportional to the column sums of the available frequencies. The unbiasedness is proved for the case of $k = 3$ columns and $l = 2$ rows only, providing the last frequency in the first row ($= f_{13}$) is missing. For the same special case asymptotic formulae for the variances and covariances of the estimates are obtained by inversion of the information matrix associated with the likelihood function.

(B. Schneider)

2/683

CHOW, G. C. (Cornell University, Ithaca, New York) **6.1** (5.2)
Tests of equality between sets of coefficients in two linear regressions—*In English*
Econometrica (1960) **28**, 591-604 (10 references)

In this paper the author presents systematic tests which are involved when, having estimated a linear regression with p coefficients, it is necessary to test whether m additional observations belong to the same regression. The author relates the prediction interval for $m = 1$ with the analysis of covariance for $m>p$ within the framework of the general linear hypothesis for any m. These results are then extended in order to test the equality between subsets of coefficients.

Firstly, the author discusses various applications of normal linear regression in the representation of an economic relationship and points out some of the complications and the methods used to overcome them. His first test concerns the hypothesis that both samples belong to the same regression and uses the well-known prediction interval, see Mood [*Introduction to the Theory of Statistics* (1950) New York: McGraw-Hill, pp. 304-305]. This can be utilised when the number m of observations in the second sample equals one; the analysis of covariance test, see Kendall, M. G. [*The Advanced Theory of Statistics* (1946) **2**, London: Griffin, p. 242], may be used when $m>p$. The author presents two tests for the case where $2\leqq m\leqq p$.

The first test is based on a prediction interval for the mean of m observations and the second test is an F-test (Section 3). Later the author explains the relationship between this F test, the prediction interval and the analysis of covariance and proceeds to test the equality between regression coefficient subsets in both the regressions. Two examples are given from the field of economic relationships concerned with the temporal stability of a statistical demand function for car ownership and of a statistical demand function for new cars. The author summarises his results as follows: "To test the equality between sets of coefficients in two linear regressions, we obtain the sum of the squares of the residuals assuming their equality, and the sum of squares without assuming their equality. The ratio of the difference between these two sums to the latter sum, after adjustment for the corresponding degrees of freedom, will be distributed as the F-ratio under the null-hypothesis. This latter sum of squares will be computed only from the first sum of n observations when the second sample is not large enough for computing a second regression."

(W. R. Buckland)

2/684

CHRIST, C. F. (University of Chicago) **6.6** (11.7)

Simultaneous equation estimation: any verdict yet?—*In English*

Econometrica (1960) **28**, 835-845 (7 references)

This paper is one of four papers presented by various authors in (1958) at the meeting of the Econometric Society. It is a revised version of the original paper and here the position taken is slightly more in favour of simultaneous equations methods than that taken previously; [see also abstracts No. **2**/688, **6.6**, No. **2**/690, **6.6** and No. **2**/692, **6.6**]. The relative merits of ordinary least-squares estimation and several simultaneous equations methods are discussed and the importance of the idea of a system of simultaneous structural economic equations is stressed. The available evidence suggests that on occasions while economical least-squares methods may give results which are as good or better than the simultaneous equations method it is preferable to use the latter.

After giving some standard definitions and remarking that Monte Carlo experiments are more often applicable as there is yet little small-sample theory of the properties of the different estimators, the author begins by discussing over-identified linear stochastic models. A model is stated and methods for obtaining its reduced form given. The estimation methods discussed for the equations of the model, termed " structural equations " are: (i) ordinary least-squares, (ii) two stage least-squares, (iii) limited information, (iv) full information maximum likelihood. Estimation methods are also proposed for a reduced form equation.

A discussion of the literature concerning the equations and their properties follows and other related evidence is commented on at some length. The evidence provided by Monte Carlo experiments is stated and various comparisons made. The author then reports on three Monte Carlo studies by Wagner [*Econometrica* (1958) **26**, 117-133]: Basman [A paper presented at the Econometric Society Meeting (1958)] and Summers [A paper presented at the Econometric Society Meeting (1958)] and comments on work by Nagar [*Econometrica* (1959) **27**, 575-595: abstracted in this journal No. **2**/300, **4.8**, and *Econometrica* (1960) **28**, 573-590: abstracted in this journal No. **2**/777, **11.7**] and by Theil [*Statistical Estimation of Simultaneous Economic Relationships* (1958) Amsterdam: North Holland Publishing Co.] and states that Nagar has done further experiments with Wagner's (1958) models. In summing-up the author says it is not yet clear that the least-squares method for structural equations is dead or should be discarded.

2/685

(W. R. Buckland)

CSÁKI, P. & FISCHER, J. (Math. Inst., Hungarian Academy of Sciences, Budapest) **6.2** (6.9)

On bivariate stochastic connection—*In English*

Publ. Math. Inst. Hung. Acad. Sci. (1960) **5**, 311-323 (8 references, 1 figure)

The bivariate stochastic connections are investigated with the aid of the theory of Hilbert spaces. Let (Ω, S, P) be a probability field, L^2 the Hilbert space of the random variables ξ with finite variance, S_ξ the smallest σ-algebra for which ξ is measurable and $L^2_\xi = L^2(\Omega, S_\xi, P)$; further

$$L^2_F = \left\{ f(x) \int_{-\infty}^{\infty} f^2 dF(x) < \infty \right\};$$

$L^2_{\xi o}$ denotes the subspace of L^2_ξ corresponding to zero expectation. Let $A_\xi \zeta$ be the orthogonal projection of ζ to L^2_ξ that is, $A_\xi \zeta = M(\zeta \mid \xi)$. Then the measures of the stochastic connection, the correlation coefficient $R(\xi_1, \xi_2)$, and the correlation ratio $\Theta_\xi(\zeta)$ may be written as $R(\zeta_1, \zeta_2) = (\zeta_1^*, \zeta_2^*)$; $\Theta_\xi(\zeta) = \| A_\xi \zeta^* \|$ where ζ^* is the standardised form of ζ. For the maximal correlation S and for the contingency C the following relations hold:

$S(\xi, \eta) = \| A_\xi \| = \| A_\eta \|$; $C(\xi, \eta) = ||| A_\xi ||| = ||| A_\eta \||$

provided that the domains of A_ξ and A_η are restricted to $L^2_{\xi o}$ and $L^2_{\eta o}$.

Let A_1 be the operator, corresponding to A_η, mapping $L^2_{H_1}$ into $L^2_{H_2}$; H_1 and H_2 are the density functions of ξ and η, respectively, and $H(x, y)$ is their joint density function. Similarly the operator A_2 corresponding to A_ξ can be defined. A_1 and A_2 are integral operators if and only if the measure generated by H is absolutely continuous with respect to the product of the measures P_1 and P_2 generated by H_1 and H_2 respectively. With the aid of a theorem, the above expression of the contingency can be proved to be equivalent to the definition given by Rényi, "New version of the probabilistic generalisation of the large sieve " [*Acta Math. Acad. Sci. Hung.* (1959) **10**, 441-451: abstracted in this journal No **1**/252, **6.0**].

Replacing the operator A_1 by $\bar{A} = A_1 A_2$ instead of the measure P a symmetric measure $\bar{P}$ is obtained. $\bar{H}(x, y)$ being the distribution function corresponding to $\bar{P}$ we have

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)\, g(y)\, d\bar{H}(x, y) = (A_2 A_1 f(x), g(x)).$$

In addition the square of the maximal correlation of $H(x, y)$ equals the maximal correlation of $H(x, y)$. Computations will be often facilitated by using this connection.

2/686

(P. Bártfai)

CSÁKI, P. & FISCHER, J. (Math. Inst., Hungarian Academy of Sciences, Budapest) **6.2** (6.0)

Contributions to the problem of maximal correlation—*In English*

Publ. Math. Inst. Hung. Acad. Sci. (1960) **5**, 325-337 (5 references, 2 figures)

Preserving the notations of authors' preceding paper, see " On bivariate stochastic connection " [*Publ. Math. Inst. Hung. Acad. Sci.* (1960) **5**, 311-323; abstracted in this journal No. **2**/686, **6.2**] let the domain of A_ξ be $L^2_{\xi 0}$, then the functions $f \in L^2_{\xi 0}$, $g \in L^2_{\eta 0}$ form a pair of eigenfunctions with eigenvalue λ, if $A_\xi g = \lambda f$ and $A_\eta f = \lambda g$. If f and g are standardised variates then the following three statements are equivalent:

(i) f and g form a pair of eigenfunctions,

(ii) $\Theta_\xi(g) = \Theta_\eta(f) = (f, g)$,

(iii) f and g are linearly correlated and $\Theta_f(g) = \Theta_\xi(g)$ $\Theta_g(f) = \Theta_\eta(f)$.

In the remaining part of Section one the case $S(\xi, \eta) = 1$ is considered and it is proved that the linearly correlated standardised variates ξ and η are homoscedastically correlated if and only if $\xi^2 - 1$ and $\eta^2 - 1$ form a pair of eigenfunctions with eigenvalue $(\xi, \eta)^2$.

The computation of the maximal correlation is facilitated for some cases if $\{\phi_k\}$ and $\{\psi_k\}$ are linearly independent systems in $L^2_{\xi 0}$ and in $L^2_{\eta 0}$, respectively, then the functions $f_n = \sum_{k=1}^{n} \alpha_{kn}\phi_k$ and $g_n = \sum_{k=1}^{n} \alpha'_{kn}\psi_k$ are eigenfunctions if and only if, $A_\eta \phi_n$ and $A_\xi \psi_n$ can be written as linear combinations of $\psi_1, \psi_2, \ldots, \psi_n$ and $\phi_1, \phi_2, \ldots, \phi_n$ respectively; the eigenvalue can easily be calculated from the coefficients. If the eigenfunctions are polynomials then the choice of $\phi_n = \xi^n - M(\xi^n)$ is suitable.

The second section of this paper gives examples for the determination of the maximal correlation.

(P. Bártfai)

2/687

HILDRETH, C. (Michigan State University) **6.6** (11.7)

Simultaneous equations: any verdict yet?—*In English*

Econometrica (1960) **28**, 846-854 (25 references)

This paper is the second in the series presented at the Symposium [see also abstracts No. **2**/685, **6.6**, No. **2**/690, **6.6** and No. **2**/692, **6.6**]. The author discusses his reasons for believing that the experience with applications of existing simultaneous equation procedures does not yet furnish a reliable proof of their practical usefulness.

He gives an example of a situation which is not yet covered by these studies and gives extensive references as to the work already completed. While deploring such statements as "Least-squares methods are obsolete" he remarks that at present such statements are unjustified and can present a biased view of the basis upon which to choose methods and models.

He discusses both changes in the type of model and work being done on a new estimation procedure [*Econometrica* (1959) **27**, 302] and stresses the importance for the future of the choice of the best procedure to employ in various circumstances. The author points out that even minor generalisations of traditional models have led to severe mathematical complications. This has lead to difficulties in the development of procedures and made almost non-existant finite sample distribution and test theory. The case in which the Monte Carlo method fills an important gap is stated generally. The importance of testing predictions is stressed and the way in which the Monte Carlo method can be used in this context is discussed.

Finally, the author remarks that he would not wish to give a verdict on simultaneous equations at this time but would like to see theoretical and empirical investigations continued.

(W. R. Buckland)

2/688

JOHNSON, E. A. & BROWN, B. W. (School of Public Health, Univ. Minnesota, USA) **6.8** (4.3)

The Spearman estimator for serial dilution assays—*In English*

Biometrics (1961) **17**, 79-88 (9 references, 1 table, 2 figures)

To estimate the density of a specified organism in a given suspension, a common method is to take *n* doses of constant volume from each of a series of successive dilutions of the original suspension and to observe the proportion of doses which show growth of the organism. The dilutions may be so chosen that the logarithms of the resulting concentrations are equally spaced, and in this case Spearman has proposed an estimator for the logarithm of the density of organisms in the original suspension which is a linear function of the original log-concentration, the difference between successive log-concentrations, and the observed proportion of doses showing growth.

Assuming a Poisson distribution of organisms per dose in the original suspension and assuming such a range of concentrations and such a dosage that it is very probable that growth will occur in a dose from the original suspension and very unlikely in a dose from the lowest concentration, the authors show that the Spearman estimator has negligible bias and find its approximate variance which is of order $1/n$. Based on these results and the fact that the estimator is asymptotically normal, confidence limits for the log-density are proposed using the normal approximation. The authors also find the approximation to the mean and variance of the estimator of the density derived from Spearman's estimator of the log-density.

Two graphs are provided to facilitate finding the point estimate and interval estimate of the density.

(C. C. Thigpen)

2/689

KLEIN, L. R. (University of Pennsylvania) **6.6** (4.1)

Single equation versus equation system methods of estimation in economics—*In English*

Econometrica (1960) **28**, 866-871 (3 references)

This paper, the fourth of the symposium (see also abstracts No. **2**/685, **4.1**, No. **2**/688, **4.1** and No. **2**/692, **4.1**) is a discussion by the author on the history of the various developments in the methods of estimation and includes his comments on various points such as bias, misuse of the choice of method to be used for estimation and the way in which the use of the computers of various sorts has influenced the field of estimation. He points out that limited-information estimates can now be obtained in a matter of minutes and high speed programmes are being made available for the full maximum-likelihood estimation. Recent interesting Monte Carlo results have also been made possible by the advent of high-speed computers.

When comparing the two methods of estimation the author remarks that a comparison coefficient by coefficient of structural estimates is often not a suitable approach and he sees the need for comparisons of summary statistics of a whole system. Reduced-form coefficients may differ widely even though component structural coefficients appear closely related. The efficiency properties of single-equation least-squares estimates do not hold under the transformation from structural to reduced-form coefficients. Finally the author discusses the applicability of least-squares.

(W. R. Buckland)

2/690

KLOEK, T. & MENNES, L. B. M. (Netherlands School of Economics) **6.6** (6.4)

Simultaneous equations estimation based on principal components of predetermined variables—*In English*

Econometrica (1960) **28**, 45-61 (7 references, 6 tables)

Estimation of economic equation systems is usually based on rather short time series: if such a system happens to be large this gives rise to difficulties in all cases where the estimation method used requires estimates of the moments of reduced-form disturbances. These difficulties often arise in cases of limited-information, maximum-likelihood and two-stage least squares estimations. In this paper the author demonstrates a practical solution of solving these difficulties by means of the use of a small number of principal components of predetermined variables which are used to represent the complete set of the predetermined variables.

In the first part the author discusses over-identification and undersized samples and shows the normally adopted procedures. An account of some experiments with the introduction of principal components of predetermined variables and of certain least-squares residuals is given. The author says that although these approaches do not avoid all disadvantages they do tackle one procedure with a certain amount of success.

The third section gives the common part of the alternative estimation procedure and the subsequent sections discuss the use of four variants. These are applied numerically in the last section of the paper.

One of the six tables used to illustrate this paper gives the alternative of Klein's model I, [Klein, L. R. & Goldberger, A. S. " An econometric model of the United States from 1929-1952 " (1955) Amsterdam].

(W. R. Buckland)

2/691

LIU, T. C. (Cornell University, Ithaca, New York) **6.6** (4.8)

Under-identification, structural estimations, and forecasting—*In English*

Econometrica (1960) **28**, 855-865 (13 references)

In a system of simultaneous linear equations a structural relationship is identifiable only if the following condition is met:

Rank $[\Pi_{\Delta,\,**}] = G^{\Delta}-1$; $K^{**} \geqq G^{\Delta}-1$ where $[\Pi_{\Delta,\,**}]$ is the matrix of the reduced-form of the regression coefficients of the jointly dependent variables included in the relationship given (G^{Δ}) on the predetermined variables which are included in the structure but excluded from the given number of relationships, K^{**}. This condition appears unimportant but is necessary for identification: it is satisfied in all existing statistical estimates concerning structural relationships. Both the original supposition underlying the simultaneous equation approach and, in the empirical models, the universal fulfillment of the necessary condition of identification, present an interesting situation. Economic variables are considered as mutually dependent but the degree of simultaneity is only recognised in so far as it does not prevent identification of the structural coefficients. Literature already available answers that we are often required to exclude a sufficient number of variables from the structural relationship either by theory or by *a priori* information. When considering economic theory, this paper suggests, one requires to include a much greater number of variables than is usual in existing structural models for economics.

The complexity of the modern economic structure makes the true structural relationship more likely to be under-identified rather than over-identified. Determined efforts to obtain " significant " structural estimates may well be the cause of the fact that all existing empirical structural relationships are over-identified. The responsibility for the over-identification as a result of over-simplification is due more to statistical difficulties than to economic theory or *a priori* information.

The author then proceeds to discuss and demonstrate the misleading nature of the type of so-called structural estimate.

Finally the paper deals with the rather doubtful prospects of over-identifying an otherwise under-identified structural relationship by resorting to more division of the time periods or by disaggregating the variables. He does not think that the results of either of these procedures are of significant use. (See also abstracts No. **2**/685, **6.6**, No. **2**/688, **6.6** and No. **2**/690, **6.6**.)

(W. R. Buckland)

2/692

MANN, H. B. (Ohio State University, Columbus, Ohio) **6.1** (9.1)
The algebra of a linear hypothesis—*In English*
Ann. Math. Statist. (1960) **31**, 1-15 (3 references)

James introduced the concept of the relationship algebra of an experimental design [*Ann. Math. Statist.* (1957) **28**, 993-1002], and demonstrated that the important properties of an experimental design, including the analysis of variance appropriate to it, can be revealed by analysing the structure of its relationship algebra. Essentially, this algebra is generated by the basic relationships that exist between the experimental units of the design.

As an illustration, James presented a detailed analysis of the relationship algebra of a balanced incomplete blocks design. Mann has considered a more generalised algebra in which the relationship algebra of an experimental design is a special case. He illustrates his methods by decomposing the algebra of an *s*-dimensional cubic lattice into its principal components.

(B. S. Pasternack)

2/693

ROBERTSON, W. H. (Sandia Corporation, Albuquerque) **6.9** (11.5)
Programming Fisher's exact method of comparing two percentages—*In English*
Technometrics (1960) **2**, 103-107 (1 reference, 1 figure)

By re-writing the expression for Fisher's exact method of comparing two percentages, a high-speed computer programme is developed for computing these exact, to a degree of decimal accuracy, probabilities. The difference in two percentages is expressed first in terms of the usual 2 by 2 table with entries: a, b, c and d where a is the min (a, b, c, d). The probability of interest is the sum of this observed percentage difference or percentage more diverse from these such as $(a-1,\ b+1,\ c+1,\ d-1)$ etc. Fisher proves that the probability of the observed results when the hypothesis of no real difference in population percentage is true is given by:

$$P_1 = \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{a!\,b!\,c!\,d!\,(a+b+c+d)!}.$$

Robertson shows that this may be written as:

$$P_1 = \prod_{x=0}^{a-1} \frac{(a+b)-x}{(a+b+c+d)-x} \cdot \prod_{y=0}^{a-1} \frac{(a+c)-y}{a-y} \cdot \prod_{z=0}^{c-1} \frac{(c+d)-z}{(b+c+d)-z}.$$

This expression lends itself to a three-loop programme given in the paper where x is increased from 0 to $a-1$, then y from 0 to $a-1$ and finally z from 0 to $c-1$. The whole programme must be repeated for more diverse percentage differences than the one observed, but one can include a preset significance level such that the programme will stop when this value is exceeded.

(C. R. Hicks)

2/694

SANDEE, J. (Central Planning Bureau, The Hague) **6.0** (6.6)
A development model for India—*In Dutch*
Statist. Neerlandica (1961) **15**, 189-210 (4 tables)

A model for the economic development of India in the 1960-70 period is described from the statistical point of view. The model contains 13 sectors, each with their specific products for which balance-equations are set up. Input-output and capital-output relations connect the sectors. Agricultural output is linked to the use of fertiliser, irrigation and advice. Consumption in 1970 is maximised within thirty-five constraints. Finally, the author discusses the effects of statistical uncertainties and possible improvements of the model.

(J. Sandee)

2/695

SIDDIQUI, M. M. (Boulder Labs., National Bureau of Standards, U.S.A.) **6.1** (5.1)
Tests for regression coefficients when errors are correlated—*In English*
Ann. Math. Statist. (1960) **31**, 929-938 (7 references)

The author is continuing some earlier work in the study of regression problems in which the residual errors are not independent, see " Covariances of least-squares estimates when residuals are correlated " [*Ann. Math. Statist.* (1958) **29**, 1251-1256; abstracted in this journal No. **1**/92, **6.1**]. In the present paper he investigates the results of using the usual test statistics for regression coefficients, derived for independent residuals, when this condition is not fulfilled. He points out that it may be possible to transform the problem into the independent case, but such a move is not always wise or feasible.

With a specific covariance matrix for the errors, namely one in which elements more than one removed from the main diagonal are negligible, and with the assumption that the residual sum-of-squares and vector of parameter estimates may be considered independent, he obtains the distributions of certain test statistics, and relates these to their common counterparts. From here it is possible for him to point out that the usual statistics may be effectively used in many areas where the residuals are correlated.

The author gives applications of the theory to the cases of testing a single sample mean, the difference between the means of two samples, and the coefficient in a linear trend and in regression on trigonometric functions. A concluding statement is given indicating the meaning of the work presented and pointing out directions for possible extension.

(D. H. Shaffer)

2/696

SCHNEEBERGER, H. (Wiesbaden) **6.1** (7.0)

Theory and applications of tests on the hypothesis of linearity in total regression equations. II—*In German*

Math. Tech. Wirtschaft (1960) **4**, 175-179

The second part of the author's paper on tests of linearity contains in the first place the demonstration of the second part of the first theorem (abstracted in this journal No. **2**/107, **6.1**). It is proved by the statement of $\mathscr{E}s_1^2 > \mathscr{E}s_2^2$ assuming $H_1 : \mathscr{E}y_{ij} = \alpha_j + \beta_j (x_{ij} - \bar{x})$, that is to say, not all the straight lines received for T_2-type groups and/or not all the means received for T_1-type groups coincide with the straight line $\mathscr{E}y = \alpha + \beta (x - \bar{x})$ generally valid on H_0.

It is shown that the standard technique to test the linearity in total regression equations follows from the first theorem in the special case of $m_1 = 0$ and therefore $m = m_2$. It is also shown that theorem 1 is a generalisation of a fundamental theorem in analysis of variance, where $H_0 : \mathscr{E}y_j = \alpha$ for each j.

In chapter two of the paper, theorem 1 is generalised to theorem 1′ by weakening the third assumption to $\mathscr{E}(y_j - \mathscr{E}y_j)^2 = \mathscr{E}(y - \mathscr{E}y)^2/g_j$; $g_j > 0$; originally in theorem 1: $\mathscr{E}(y - \mathscr{E}y)^2 = \text{const}$ for each x. Now, in the statistic F of theorem 1, s_1^2 is to be replaced by

$$(s_1')^2 = \sum_j g_j \sum_i (y_{ij}^{*} - y_{ij}^{**})^2/(m_1 + 2m_2 - 2) \text{ and } s_2^2 \text{ by}$$

$$(s_2')^2 = \sum_j g_j \sum_i (y_{ij} - y_{ij}^{**})^2/(n - m_1 - 2m_2).$$

All other conditions and results remain. Theorem 1′ is a very theoretical one, for the g_j's will never be known in practice, and their estimates will always be biased.

As a by-product of theorem 1′ the author obtains a second theorem to test the identity of m regression coefficients α_j and β_j simultaneously. He interprets the $m = m_2$ T_2-type groups ($m_1 = 0$) as stochastically independent sets of x values and tests the simultaneous hypothesis H_0: $\alpha_j = \alpha$ and $\beta_j = \beta$ for each j by $F = (s_1'')^2/(s_2'')^2$ with $(s_1'')^2 = \sum_j g_j \sum_i (y_{ij}^{*} - y_{ij}^{**})^2/2(m-1)$ and $(s_2'')^2 = \sum_j g_j \sum_i (y_{ij} - y_{ij}^{**})^2/(n - 2m)$, F-distributed with $f_1 = 2(m-1)$ and $f_2 = n - 2m$ degrees of freedom.

continued

SCHNEEBERGER, H. (Wiesbaden) **6.1** (7.0)

continued

Theory and applications of tests on the hypothesis of linearity in total regression equations. II—*In German*

Math. Tech. Wirtschaft (1960) **4**, 175-179

If H_0 cannot be rejected, central confidence intervals may be specified for α and β by $\alpha = a \pm ts/\sqrt{(\sum_i g_j n_j)}$

and $$\beta = b + ts/\sqrt{\{\sum_j g_j \sum_i (x_{ij} - \bar{x})^2\}},$$

with

$$\bar{x} = \sum_j g_j \sum x_{ij}/\sum_j g_j n_j \text{ and } s^2 = \sum_j g_j \sum_i (y_{ij} - y_{ij}^{*})^2.$$

The author states that he intends to publish a further generalisation of his theorems in which a single independent variable x will be replaced by a set of k independent variables $x_1, x_2, \ldots, x_k$.

(R. K. Bauer)

ABT, K. (U.S. Naval Weapons Lab., Dahlgren, Virginia) **7.6** (6.1)

Analysis of covariance and analysis of differences. II—*In French*

Metrika (1960) **3**, 177-211 (7 references, 1 table, 3 figures)

The first part of this study, see [*Metrika* (1960) **3**, 26-45 and 95-116: abstracted in this journal No. **2**/111, **7.6**] dealt with the analysis of covariance with one independent variable. The second part discusses the relations between the analysis of covariance and the analysis of variance of differences.

Referring to linear regression analysis it is first shown that the generalised differences $d = y-\beta_0 x$, where β_0 is postulated, can be properly used for measuring the increment, adjusted if $\beta_0 \neq 1$, of a variable if y is linearly dependent on x and if β_0 equals the true value of the regression coefficient of y on x. Here x and y, respectively, stand for the values of the variable before and after a " treatment " being applied to the experimental unit on which d is measured or observed. The test of the null hypothesis $\mathscr{E}(b) = \beta_0$, therefore, is referred to as " Test for Admissibility of Differences ".

Then, for the case of t groups of paired observations " before and after treatment ", the power of the test for the null hypothesis " no group effects " in the one-way analysis of variance of generalised differences is found to be at least equal to that of the analysis of covariance if again $\mathscr{E}(b) = \beta_0$; b this time being the regression coefficient for error " within-groups ". The F-test in the analysis of differences is even more powerful if besides $\mathscr{E}(b) = \beta_0$ appreciable differences between the group means in x exist. This applies especially to the case of $t = 2$ groups.

For computational purposes the F values of the two types of analyses themselves are compared. For the special but most important case of " ordinary " differences ($\beta_0 = 1$) the analysis of covariance can be bypassed if in fact the F-test in the comparatively simple analysis of differences is of equal or higher power as can be concluded from certain procedures described in the paper. Numerical examples are given to illustrate the findings.

The results found for the one-way classification are then generalised for the cases of two and more ways of classification of the data.

(K. Abt)

2/699

BHAPKAR, V. T. (University of North Carolina, Chapel Hill) **7.1** (9.1)

Confidence bounds connected with ANOVA and MANOVA for balanced and partially balanced incomplete block designs—*In English*

Ann. Math. Statist. (1960) **31**, 741-748 (5 references)

Simultaneous confidence bounds on particular sets of parametric functions associated with a strongly testable linear hypothesis for either ANOVA (analysis of variance—univariate) or MANOVA (multivariate analysis of variance) have been obtained by S. N. Roy & J. Roy [*Inst. Statist.*, Univ. North Carolina Mimeo. Series No. 208, (1958)]. These parametric functions can be generally regarded as measures of deviation from the " total " hypothesis and its various components, and are such that for any specific problem one of them can be appropriately designated the " total " parametric function while the others are " partial " parametric functions of various orders. The " total " function, firstly in the univariate case, is related to the non-centrality parameter of the usual F-test of the " total " hypothesis in ANOVA, and secondly, in the multivariate case, is the largest characteristic root of a certain parametric matrix which is related to another parametric matrix whose non-zero characteristic roots occur as a set of non-centrality parameters in the power function for the test, regardless of which of the standard tests is used, of the total hypothesis in MANOVA. This statement also applies to " partial " functions of various orders and their related " partial " hypothesis.

In this paper for both ANOVA and MANOVA, the author considers the hypothesis of equality of treatment effects and vector equality in the multivariate case. He then proceeds to develop explicit algebraic expressions for the total and partial parametric functions that are attached to the simultaneous confidence statements subsumed under the hypothesis of equality of treatment effects, both for ANOVA and MANOVA. The author later indicates the procedure for obtaining algebraic expressions for the confidence bounds on each such parametric function, without a derivation of these expressions in an explicit form.

(B. S. Pasternack)

2/700

COX, D. R. (Birkbeck College, London) **8.8** (1.8)
Serial sampling acceptance schemes derived from Bayes' theorem—*In English*
Technometrics (1960) **2**, 353-360 (4 references, 3 tables)

Since consecutive batches of a sampling scheme may be correlated, the author proposes an acceptance scheme wherein one examines adjacent samples before sentencing a " lot ". A stochastic process is set up representing the system and Bayes theorem is applied to obtain a sentencing rule.

The model is as follows: consider batches ..., B_{n-1}, B_n, B_{n+1}, ..., with counted defectives in random sample ..., X_{n-1}, X_n, X_{n+1}, ...; let $\mathscr{E}(X_n) = m_n$. The assumption that m_n either has the value a or b $(a<b)$ is made, and if $m_n = a$ the batch is good; if $m_n = b$ the batch is bad. Runs of good and bad batches are distributed independently with mean run length $1/t_a$ and $1/t_b$. It follows that the prior odds that a batch is bad is t_a/t_b: if $t_a+t_b = 1$ then the sequence is truly random. Considering only B_{n-1} and B_n;

$$\Pr[m_{n-1} = m_n = a] = t_b(1-t_a)/(t_a+t_b),$$

$$\Pr[m_{n-1} = m_n = b] = t_a(1-t_b)/(t_a+t_b),$$

$$\Pr[m_{n-1} = a,\ m_n = b] = \Pr[m_{n-1} = b,\ m_n = a] = \frac{t_a t_b}{t_a+t_b}.$$

If before sentencing B_n we consider $(X_{n-k}, X_{n-k+1}, \ldots, X_n)$ the author's rule is to accept B_n if, in view of the above probabilities and the assumed prior distribution $\Pr[m_n = a]$ is sufficiently large. If W_b and W_a are respective losses from accepting a bad batch and rejecting a good, then we accept if $\Pr[m_n = a_n] < W_b/W_a$. Each scheme is then characterised by a rejection rule, an expected loss, and by probability of an incorrect decision given a prior distribution.

(E. H. Lehman, Jr.)

2/701

DAVIES, O. L. (I.C.I., Macclesfield, England) **8.8** (8.7)
Some statistical aspects of the economics of analytical testing—*In English*
Technometrics (1959) **1**, 49-61 (6 references, 3 figures)

General methods are developed for determining the number of analytical tests, n, to be made on a batch of chemical products in order to minimise the total cost, testing cost plus the cost due to losses by incorrect decisions on batches. The author extends and applies the methods of others to routine laboratory testing of batch products. The assumption of a linear increasing testing cost and quadratically decreasing losses is made throughout.

Two situations are explained: firstly, when the quality of a product is measured on a continuous scale the problem is to find a point on this scale, actually on the scale of averages based on n repeat tests, where the risk of passing B grade (bad) material is fixed and then the total cost is minimised. If T is the total cost per batch, the author minimises $T(=Pc+nb)$ subject to the restriction that $p(-a;\ X, n) = \alpha$, where P is the proportion of A grade (good) material rejected by the plan, c is the cost per batch for reprocessing, b is the cost per test of testing, X is the critical point beyond which a batch is rejected, and $-a$ is the quality which should be rejected. An example is given where the distribution of quality and errors are both assumed to be Gaussian: the proportion of " good " batches rejected is determined from bivariate normal tables, as quality and errors are both assumed normal.

The second situation discussed is the case where the amount of intermediate product to be added " in process " for maintaining a given final quality is tested. The same problem of optimisation is posed; but the loss curve is assumed to be a quadratic function of x, where x is the per cent. error from the optimum amount that should be added: or, $L = ax^2$. In the appendix it is proved that the average loss due to errors in the amount added $= a\sigma_0^2/n$, where σ_0 is the standard error of testing. Thus the total cost per batch is: $T = (a\sigma_0^2/n)+nc$, here c is the cost per test. Minimising, we find $n = \sigma_0\sqrt{(a/c)}$. If prior information is available on batch to batch variation (σ_1), he shows that

$$T = nc + a\,\frac{\sigma_0^2\sigma_1^2}{\sigma_0^2+n\sigma_1^2}$$

and for a minimum:

$$n = \sigma_0\sqrt{\left(\frac{a}{c}\right)} - \frac{\sigma_0^2}{\sigma_1^2},$$

which is a smaller number of tests than when we have no prior information. (C. R. Hicks)

2/702

GOLDSMITH, P. L. & WHITFIELD, H. (Imperial College, London) **8.9** (5.7)

Average run lengths in cumulative sum chart quality control schemes—*In English*

Technometrics (1961) **3**, 11-20 (5 references, 2 tables, 8 figures)

This paper's primary aim is to give curves for the average run length before action for cumulative sum chart schemes. The cumulative sum is the algebraic sum of the deviations from some desired process mean, which is plotted against time. A V-mask with angle $2\,\theta$ is placed with vertex d time-units to the right of the last point and with axis horizontal. If any points lie outside the V, action is taken.

Two process levels μ_0 when a process is on target, and $\mu_1 = \mu_0 + k\sigma$ are to be specified, then the average run lengths are to be specified for each, perhaps $L_0 = 500$ (X's), and some small amount L_1 for μ_1.

The average run lengths were carried out by Monte Carlo simulation for $k = 0, 0{\cdot}5, 1, 2, 3$ with d's 1, 2, 5 and 8 and suitable values of θ. A comparison is then made between the present cumulative sum procedures and Shewhart charts for individual X's, using only limits at $\pm 3{\cdot}09\,\sigma$, without warning limits and runs. Average run lengths for cumulative sum charts are considerably smaller for low k's than for this particular X-chart usage. Empirical formulae are given for L_0 and L_1 in terms of d and θ. Curves are also given for the average run length in various cases where there is serial correlation in the data.

(I. W. Burr)

2/703

HALPERIN, M. & BURROWS, G. L. (Knolls Atomic Power Laboratory) **8.8** (1.6)

An asymptotic distribution for an occupancy problem with statistical applications—*In English*

Technometrics (1961) **3**, 79-89 (1 reference, 3 tables, 1 figure)

In this paper the authors consider a problem of N items of which D are defective. Many times the items are divided into k lots of equal size s, and it is desired that the number of defectives in each lot be less than a given number, x. It is shown that the number of lots of x defectives or more, $m\,(x)$, is asymptotically multivariate normal. Thus $\Pr\,\{m\,(x) \leqq m_0\}$ can be determined.

The asymptotic results are derived through an occupancy argument, and the adequacy of the asymptotic theory is discussed. Applications to acceptance sampling are indicated.

(P. H. Randolph)

2/704

HALD, A. (University of Copenhagen) **8.8** (2.4)

The compound hypergeometric distribution and a system of single sampling inspection plans based on prior distributions and costs—*In English*

Technometrics (1960) **2**, 275-340 (24 references, 18 tables, 12 figures)

This long and full paper is divided into fourteen sections. Section one is introductory. Sections two to five define and discuss conventional terms " Operating Characteristic " and " Acceptable Quality Level " and describe and criticise four Quality Control systems, namely the Statistical Research Group of Columbia, U.S. Military Standard 105, Philips Standard Sampling and Dodge-Romig. The author's criticism on all these systems is the arbitrariness of certain assumptions, for instance, always choosing the producer's risk to be some fixed percentage, ignoring cost of accepted defectives, or not precisely defining the prior distribution.

Section six presents two models based on cost minimisation: the average cost per item submitted in model one consists of three terms:

(1) sampling inspection cost,
(2) expected loss due to accepted defectives,
(3) cost of a rejected lot times probability of rejection.

This model does not consider manufacturing costs: in model two an adjustment is made to include manufacturing costs.

Section seven derives the mean, variance, and covariance of the two variables x, the number of defectives in the sample, and $y = X-x$, the remaining number of defectives in the uninspected portion of the lot; given that x is hypergeometric with parameters N, the lot size, n the sample size, and X the number of defectives in the lot: and that X itself is subject to some prior density function $f_N(X)$.

The formulae which are lengthy for σ_x^2, σ_y^2 and σ_{xy} are given, depending upon the original $f_N(X)$. This density function is called normal, hypernormal or subnormal as $\delta_N = 0, >0, <0$ in this formula:

$$\sigma_X^2 = N\bar{p}\bar{q}\,(1+\delta_N),\ \delta_N \geqq -1, \text{ where } \bar{p} = 1-\bar{q} = \mathscr{E}(X)/N.$$

Section eight shows how if $f_N(X)$ is hypergeometric, binomial, Polyá, or any weighted average of these where the weights do not depend upon N and X, then x has the same distribution as X with n substituted for N. That is, $f_N(X)$ is " reproducible " by random sample selection.

Section nine defines the cost $K(n, c)$ of a single

continued

HALD, A. (University of Copenhagen) **8.7** (2.4)

continued

The compound hypergeometric distribution and a system of single sampling inspection plans based on prior distributions and costs—*In English*

Technometrics (1960) **2**, 275-340 (24 references, 18 tables, 12 figures)

sampling plan as a function of n the sample size, and c the acceptance number.

$$K(n, c) = n(k_s-k_r)+Nk_r + (N-n)\sum_{x=0}^{c}[p_n(x \mid X)-k_r]\,g_n(x)$$

where k_s = cost of inspection per item, k_r = cost per item of rejected lots, $p_n(x \mid X)$ = binomial probability of exactly x defectives in a sample of size n given there are X defectives in lot of size N, and $g_n(x)$ = marginal density function of x =

$$g_n(x) = \sum_{X=0}^{N} f_N(X)\,p_n(x \mid X).$$

An optimum sampling plan chooses n and c to minimise K given the other parameters. This involves study of several secondarily defined parameters such as the average fraction defective in the non-inspected part of accepted lots

$$\bar{p}_n(c) = \sum_{x=0}^{c} p_n(x)\ g_n(x) \Big/ \sum_{x=0}^{c} g_n(x),$$

and the quantity

$$\gamma(n, c) = \sum_{x=0}^{c} g_n(x)\,[k_r-p_n(x)].$$

The optimum plan is shown to be the solution with respect to n and c of the two inequalities:

$$p_n(c) \leqq k_r \leqq p_n(c+1)$$

and

$$(N-n-1)[\gamma(n+1, c)-\gamma(n, c)] < k_s-k_r+\gamma(n, c) \leqq (N-n+1)[\gamma(n, c)-\gamma(n-1, c)].$$

Exact solutions depend upon the prior distribution $f_N(X)$ of X the number of defectives in a lot of size N.

The following three sections give optimal solution for $K(n, c)$ in the cases where $f_N(X)$ is rectangular, Polyá, and binomial respectively. The Polyá density function is similar to the binomial except that p is not constant: instead $p = p(V, \mu, \gamma) = (p+V\gamma)/[1+(V+\mu)\gamma]$ where V = number of defectives at any time, μ = number of effectives at that time, and γ is a parameter which may have any real value. If it is zero, the density function is binomial.

Tables are given to show optimum n, c, and $K(n, c)$ for various forms of the prior distribution and for various values of the parameters. The paper concludes with a summary and a bibliography.

(E. H. Lehman, Jr.)

JAMBUNATHAN, M. V. (University of Mysore, India) **8.3** (4.3)
A note on the efficiency of double sampling for stratification—*In English*
Sankhyā (1960) **22**, 367-370 (3 references)

To estimate the total value of a character y in a population double sampling may be employed if an auxiliary character x, which can be more economically measured, is available. In this method a large sample, called the first-stage sample, serves as a basis for stratification of the population according to the value of x, and at the second stage, stratified sampling is adopted for measuring the main character y and hence estimating the population total of y.

Neyman, who introduced this method [*J. Amer. Statist. Ass.* (1938) **33**, 101-116] also considered the problem of optimum allocation of the sampling units to the large and small samples, under the assumption that the cost of a single observation on y is the same in all strata.

In an earlier book [*Some aspects of sampling* (1953) Bangalore City: India Book Company] the author considered the more general case of the cost of single measurement varying from stratum to stratum and obtained approximate expressions for optimum allocation. In this note he obtains an expression for the variance of the estimate under that optimum allocation. He compares it with the variance in the case of a simple random sample and obtains a condition which causes the double-sampling method to be more efficient.

(T. V. Hanumantha Rao)

2/707

MARTIN, H. (Deutsch. Akad. Wissenschaften, Berlin) **8.8** (1.8)
Probability of misclassifications. I. Variables—*In German*
Qualitätskontr. Operat. Res. (1960) **5**, 109-112 (4 references, 5 figures)

Given a variable x, a population, for example an inspection lot, may be classified into three classes of quality in the following way: the population shall be of type A if $\mu<x_1$, B if $x_1\leqq\mu<x_2$, C if $\mu\geqq x_2$. If the sample result $\bar{x}\pm ts/Nn$ contains the boundary x_1 or x_2, to which class should the population be attached? Here n denotes the sample size, $\bar{x}$ the sample mean, s the sample-standard deviation, and t is the factor according to the level of significance chosen.

Theoretically, the problem may be solved in a very simple manner by increasing n to $n_i = [ts/(\bar{x}-x_i)]^2$; $i = 1, 2$, respectively. Practically, this method is impossible if the unknown μ lies near x_1 or x_2. The alternative method of solution is to classify by the original sample and to tolerate a certain probability of misclassifications (p_{mis}): often, a compromise between the theoretical and the practical way may be chosen.

Under the assumption of normality the author calculates the probability of misclassification by a Bayesian argument but without stating this. A diagram shows this probability for different sample sizes and position of the sample mean. The assumption of uniform distribution of the true means is used for calculating the mean p_{mis} dependent on the fraction $\frac{(x_2-x_1)\sqrt{n}}{s}$ where x_2 and x_1 are boundaries of the classes of quality. With reference to the mean p_{mis} the author discusses the possibilities of determining the necessary sample size for a given maximum of p_{mis}, and of improving a given classification scheme. He also refers to the consequences which may result from classifying in a lower category in order to be on the safe side.

(R. K. Bauer)

2/708

OLDS, E. G. (Carnegie Institute of Technology, Pittsburgh) **8.9** (5.1)
Power characteristics of control charts—*In English*
Annual Conv. Trans., Amer. Soc. Qual. Contr. (1960) 195-208 (20 references, 6 tables)

This article is a summary of several papers on power characteristics of control charts. Numerous statements are abstracted from these papers. Segments of some tables from these papers are presented to illustrate the effects on power of the number of samples, sample size, non-centrality, and width of the control limits. The author considers the case where a previous standard is not given. Hence, it is necessary to collect several samples, estimate the mean and standard deviation, and compute the control limits. The process is then defined as being in a state of control if all of the sample points fall within the control limits. A process is in a state of control during a time interval if all of the x's are independent and $f(x)$ is identical for every x.

For counts, the control limits are $\bar{c}+a\sqrt{\bar{c}}$ where $\bar{c}$ is the average number of counts from k samples. The probability of rejecting the null-hypothesis of a state of control when in fact the process is in a state of control is α. In general, exact computations of α are lengthy. An upper bound, α_u, for α was found such that $\alpha_u \rightarrow \alpha$ as $k \rightarrow \infty$. Two tables of α_u are presented which show the effects of k, a, and the true average number of counts.

A single slippage in a process occurs when during all k sub-intervals the true means are equal to c_1' except for a single sub-interval with c_2'. The probability that at least the sample point in the slippage sub-interval is out of control (probability of detecting the single slippage) was studied in relation to k, a, c_1', and c_2'. A few values are given for the power of detecting a double slippage, two intervals with true means of c_2'.

For $\bar{x}$-charts, the components of variance model, $x_{ij} = \mu_i + \varepsilon_{ij}$, was considered where μ_i is normally and independently distributed $(0, \theta^2\sigma^2)$, ε_{ij} is similarly distributed $(0, \sigma^2)$, and ε_{ij} and μ_i are independent. Control limits are determined by $\bar{x} \pm aA_2\bar{r}/3$ where x is the overall mean and $\bar{r}$ is the average range or by $\bar{x} \pm as/\sqrt{n}$ where s is the standard deviation and n is the number of samples per sub-interval. Type I and II errors were investigated for the null hypothesis that $\theta = 0$ against the alternative that $\theta > 0$ for various values of k, n, θ, and a. The analysis of variance test is more powerful than the control chart test for these alternatives. The power of the control chart test was compared with the analysis of variance test for other alternatives.

(D. W. Gaylor)

2/709

PAGE, E. S. (University of Durham Computing Laboratory, Newcastle) **8.9** (5.7)
Cumulative sum charts—*In English*
Technometrics (1961) **3**, 1-9 (21 references)

This paper traces the development of process inspection schemes from Shewhart's original work to the present techniques using cumulative sums. This includes control limits, warning limits, analysis of errors of both kinds and attendant costs, the study of average length of run before action is taken, and the extension suggested by sequential analysis to cumulative sums. The variable plotted on a cumulative sum chart is the algebraic sum of the deviations from some desired or typical process mean. Then, subjectively, if the points show a tendency down or up, that is, a change of slope, a process change is indicated.

Objective criteria are briefly discussed:

(i) take action whenever the points climb to a height of h above the lowest point or conversely fall to h below the minimum, and

(ii) placing a V-shaped mask with angle 2θ, with vertex at d time units ahead of the last point and axis horizontal, opening to the left, then take action if any points lie outside the V.

It is suggested that criteria for determining h or θ and d involve average run length before action is indicated, with long runs if process is satisfactory and short runs if " off " by a specified amount. Comparisons of average run length with Shewhart charts are made for charts for individual X's. The approach requires knowledge of process σ and stability of σ. Further problems needing solution are suggested.

(I. W. Burr)

2/710

PLESZCZYŃSKA, Elzbieta (Inst. Math., Polish Academy of Sciences, Warsaw) **8.8** (6.2)

Screening in statistical testing for correlated characteristics—*In Polish*

Zast. Mat. (1960) **5**, 47-59 (2 tables, 8 figures)

In this paper the authoress considers lots of items whose quality depends decisively on a characteristic Y, which is difficult to test. It is known, however, that Y is correlated with another characteristic X, not so important but easy to test. A further assumption is made that the pair of characteristics X, Y has two-dimensional normal distribution with known correlation coefficient ρ and known standard deviations, but with unknown means. To be concrete in subsequent description let it be assumed moreover that $\rho<0$ and that " good " items are those with Y less than a given number y_0.

In reference to an earlier discussion by Oderfeld [" Statistical testing for correlated characteristics " *Zastosowania Mat.* (1959) **4**, 255-264: abstracted in this journal No. 463, **8.8**], the authoress investigates the statistical effects of screening, that is to say the effect of rejecting items with X smaller than a given x_0 and computes the dependence between per cent. defective in a lot before and after screening.

Two acceptance procedures involving screening are discussed, the one, where the lot is rejected or screened before acceptance according to the result of estimating mean value of Y on the base of a small pilot sample, and the second, where the parameters of screening depend on the result of that pilot sample.

(S. Zubrzycki)

2/711

RAJA RAO, B. (Poona University) **8.6** (4.1)

On the relative efficiencies of BAN estimates based on doubly truncated and censored samples—*In English*

Proc. Nat. Inst. Sci. India (1958) **24**, 366-376 (6 references)

The author supposes that in drawing samples from a statistical population distributed according to a probability law $f(x;\theta)$, observations which are less than x_0 or greater than x_1 are not measured; but, the number of observations of each kind is recorded. Such a sample, we should expect, would supply less information regarding the parameter θ than a sample where all individuals in the sample are measured. The author establishes a certain inequality and states that it proves the result mentioned above. Cases where the number of observations with values less than x_0 or greater than x_1 are not separately recorded are also considered. The results are extended to cover distribution laws with more than one parameter.

(S. John)

2/712

SEN, P. K. (Calcutta University) **8.4** (4.3)

On the estimation of the population size by capture-recapture method—*In English*

Bull. Calcutta Statist. Ass. (1960) **9**, 93-110 (10 references)

In this paper some methods of estimating population size by " capture-recapture " method are discussed. A sample of size n_1 is drawn out of an unknown N and the units are " released " back into the population with a suitable marking made on each of them. Next a sample of size n is drawn and the number r of them bearing the marks previously made is noted. The problem posed by the author is to estimate N.

He considers two methods of drawing the second sample of size n: firstly by direct sampling, drawing a fixed number n of units; and secondly by inverse sampling, drawing units until a fixed number n of previously marked units appear. The cases where sampling is with replacement and also without replacement are both discussed. Since the maximum-likelihood estimators do not possess moments of any positive order in the case of direct sampling, whether it be with or without replacement, modified estimators due to Bailey [*Biometrika* (1951) **38**, 293-306] are considered. It is shown that Bailey's estimators have (1) bias which decreases more rapidly than that of the maximum-likelihood estimator, (2) finite unconditional moments of all orders, and (3) for sampling without replacement, conditional expected squared loss is not greater than that of the maximum likelihood estimator; the condition being that r is greater than a fixed positive number c. On the other hand, for sampling with replacement, the variance of Bailey's estimator is larger than that of the maximum-likelihood estimator, though both tend to the information limit as the sample size increases.

In the case of inverse sampling, the maximum-likelihood estimator is the minimum-variance-unbiased estimator when sampling is with replacement. For sampling without replacement, the maximum-likelihood estimator is biased, while Bailey's estimator is unbiased, and has a smaller variance than that of the above estimator; though the two variances are equal, asymptotically. Next, sequential procedures are discussed. A comparison of the information, in Fisher's sense, for the sequential case with that of nonsequential case shows the sequential procedure to be superior.

(T. V. Hanumantha Rao)

2/713

SEVERO, N. C. & NEWMAN, A. E. (Statist. Eng. Lab., Nat. Bur. Stand., Washington) **8.1** (4.9)

A statistical chain-ratio method for estimating relative volumes of mail to given destinations—*In English*

J. Res., Nat. Bur. Stand. C (1960) **64**, 37-47 (9 references, 4 tables, 4 figures)

A sampling method, called the " chain ratio " method, is applied in estimating the distribution of mail by destination. The details and results of the application of this sampling method to outgoing first-class letter mail in a study conducted in San Francisco are given: Los Angeles and Baltimore are other locations studied.

The basic idea involved in the estimation formulae consists of multiplying together a chain of ratios. Thus, if the records at the post offices are available, or sampling plans are devised, to give the following unbiased estimates

(i) the ratio of primary (the first stage of distribution of outgoing mail) to the total volume (T_p/T), and

(ii) the ratio of mail to the primary destination to the primary volume, (D_p/T_p),

then the product of these two ratios gives the estimate of the ratio of mail to a primary destination to the total volume, (D_p/T), or

$$(D_p/T) = (D_p/T_p)(T_p/T).$$

This concept, with modifications, is extended to obtain estimates of the ratios to the final destinations through the primary, secondary, and tertiary stages of distributions.

Variances and coefficients of variations for the estimators are given. The method of collecting data is also described.

(H. H. Ku)

2/714

STRAUCH, H. (Nürnberg) **8.9** (8.8)

Continuous sampling inspection from the belt—*In German*

Qualitätskontr. Operat. Res. (1961) **6**, 21-25 (3 references, 3 tables, 9 figures)

The problem of a combined sample and systematic inspection technique by attributes, suitable for belts or other continuous productions, has been solved theoretically by Dodge [*Ann. Math. Statist.* (1943) **14**, 264-279]. The defective units on the belt are assumed to be randomly distributed. To start the inspection, piece by piece is to be tested until i successive units are found to be acceptable; type I part of the series. From now on only every $1/f$ th unit is to be tested up to the next defective one; type II part of the series. Then another type I part of the series starts and so on. All detected defective units are replaced by non-defective ones.

Let u be the mean length of an equal number of units of type I sections and let v be the mean length of type II sections. Then the mean proportion of all units to be tested is $F = (u+fv)/(u+v)$. Furthermore, let p be the proportion of defectives in the production. After the inspection, the mean proportion of defectives is given by $D = p(1-F)$. The author calculates $u = (1-q^i)/pq^i$ and $v = 1/fp$ with $q = 1-p$. D assumes a maximum $D_{\max}$. A special sample design will be obtained by the following procedure. A certain value of $D_{\max}$ is preassigned. Then $D_{\max}$ is assumed for $p(D_{\max}) = (1+iD_{\max})/(i+1)$ and it holds

$$f = q(D_{\max})^{i+1}/\{iD_{\max}+q(D_{\max})^{i+1}\}.$$

By these formulae numerous pairs of values i and f are obtained which solve the problem. A graph given by Dodge shortens this procedure.

Another numerical example shows that u is decreasing with decreasing i and increasing f. Therefore an approximate value for p, and the efficiency of the inspection team needs to be estimated before a maximum i and a minimum f will be chosen. This and some other results in sample design follow from a broad discussion of the formulae given. Finally the author refers to some generalisations of the inspection technique proposed for cases of more than one kind of defect and of grouped lots.

(R. K. Bauer)

2/715

SUKHATME, B. V. & KOSHAL, R. S. (Inst. Agric. Res. Statist., Delhi & FAO, Cairo) **8.3** (4.1)

A contribution to double sampling—*In English*

J. Indian Soc. Agric. Statist. (1959) **11**, 128-144 (2 references)

It is well known that the efficiency of an estimate of the population total of a characteristic Y can be increased, for a given cost, if information regarding an auxiliary variable X is available or can be ascertained more economically than Y. One such technique is double sampling, where values of X are obtained for each individual in the sample, but values of Y are obtained only for individuals in a sub-sample. From the subsample, the regression of Y on X or the ratio of total Y to total X is first estimated. The estimate of total X from the whole sample is obtained and used in conjunction with the regression or ratio based on the subsample to estimate total Y.

Cochran [*Sampling Techniques* (1953) New York: Wiley] gave the formulae for the estimator, its variance and an estimator of that variance for the case of single-stage double sampling. This paper extends these results to the case of multi-stage designs and then to stratified multi-stage designs. As in the case of single-stage sampling, the estimators given are biased. A direct analogue of the conditions under which the auxiliary variable X increases the precision of the estimates is derived.

A numerical illustration is given. The authors state that they have extended these results to the case when the selection probabilities are unequal, and propose to publish them in a subsequent paper.

(T. V. Hanumantha Rao)

2/716

TIKKIWAL, B. D. (Karnatak University, Dharwar, India) **8.3** (4.1)

An examination of the effect of matched sampling on the efficiency of estimators in the theory of successive sampling—*In English*

J. Indian Soc. Agric. Statist. (1958) **10**, 16-21 (8 references, 1 table)

It was shown by the author in an earlier paper that when one or more characters of a population are studied on successive occasions by a scheme of matched sampling, the efficiency of estimators of the time-conditioned parameters on any occasion can be increased by a scheme of matched sampling. The author shows in this paper that this efficiency goes on increasing with the occasion till it reaches a limiting value.

The proof is given in two steps. Firstly, it is noted that the problem imposes certain restrictions on the sample size, etc., which reduces the multivariate case to the univariate one. Secondly it is observed that the variance of an estimator on any occasion is a certain multiple of the weight associated with the estimator on that occasion. It is shown that the weights, which are non-increasing with the occasion, tend to a limit as occasions increase. The author further examines, theoretically as well as by tabulation, the values of the weights for different values of the correlation between consecutive occasions and of the replacement fraction; also the rapidity with which the weights converge to a limiting value. It is noted that the rate of convergence decreases with the increase in absolute value of the correlation between two consecutive occasions. While proving the non-increasing behaviour of the weights in the second step, the author obtains a lower limit to the weights which provides an easy computational check on the weights.

(S. G. Prabhu Ajgaonkar)

2/717

WETHERILL, G. B. (Birkbeck College, London) **8.8** (1.8)

Some remarks on the Bayesian solution of the single sample inspection plan—*In English*

Technometrics (1960) **2**, 341-352 (2 references, 4 tables, 1 figure)

A prior distribution of the number X of defectives among K batches of size N_i is assumed binomial $b(X;\ N_i, p)$ where p has the value p_i with probability a_i; $\sum_{i=1}^{k} a_i = 1$, $p_1 > p_2 \ldots > p_k$. Let W_{1i} be the cost of accepting a batch when p_i is true and W_{2i} be the cost of rejecting a batch when p_i is true. Let p_0 be the value of p at the breakeven point when $W_{1i} = W_{2i} = 0$: then $W_{1j} = 0$ if $p_j \leqq p_0$ and $W_{2j} = 0$ if $p_j \geqq p_0$. Equations for an optimum solution take a simple form if we assume $p_i/q_i = (p')^{i/\alpha}$ where p' and α are suitable positive constants.

To optimise n and c the "neutral line" is first determined; that is, the locus of points (n, c) such that $W_{1i} = W_{2i}$. This line has the equation

$$\sum_{i=1}^{k} a_i \binom{n}{c} p_i^c q_i^{n-c} (W_{2i} - W_{1i}) = 0.$$

Next we let c be a function of n, determined by the neutral line and proceed to choose an optimal n. Let the expected loss for a sample of size n be $R(n)$. Then we seek an n such that $\Delta R = R(n+1) - R(n)$ changes from negative to positive at the optimum n. It is shown that

$$\Delta R = \tfrac{1}{2} \sum_{i=1}^{k} a_i \binom{n}{c} p_i^c q_i^{n-c} (W_{2i} - W_{1i})(2p_c - 1) + 1.$$

In order to find the optimal n one must assume the a's and W's known and then perform repeated calculations of several fairly lengthy equations, using several reasonably guessed values of n. It is shown that the optimum is "flat" and hence quite large changes in some of the parameters have little effect on $R(n)$. That is, the test is robust. The most critical parameter is p_0, the breakeven quality. Some numerical examples are given.

(E. H. Lehman, Jr.)

2/718

ZUBRZYCKA, Ludmilla (Inst. Math., Polish Academy of Sciences, Wrocław) **8.0** (1.4)

On the distribution of sampling points in a plane—*In Polish*

Zastosowania Mat. (1960) **5**, 161-171 (3 tables, 3 figures)

If we want to estimate the integral of a certain function $f(p)$ over a given region G by observing the values of $f(p)$ in n points selected from G, then it is advisable to distribute the n points as uniformly as possible over G; for example by choosing them as vertices of a triangular net. The uniformity of such a sample of points is then disturbed in a crude manner if we add a few new points to our sample. It is not so if we select the points at random; but then we must expect random clustering of points and this is not desirable.

Mrs Zubrzycka tries to find a compromise between the competitive requirements of uniformity of a sample for a given n and the preservation of this uniformity if a sample is enlarged point by point in that she defines a sequence of points p_n in the unit-square such that each segment of this sequence is distributed more uniformly than, on the average, the same number of random points. As it is known, the sequence of rests modulo 1 of nz, denoted below by t_n, where

$$z = (\sqrt{5}-1)/2 = 0{\cdot}61803\ldots$$

has remarkably good equi-partition properties over the unit interval $0 \leqq t \leqq 1$.

Advantage is taken of this by defining p_n as the image of t_n by a continuous Peano curve transforming a closed unit-interval into a closed unit-square. In table I the co-ordinates of points p_n, $n = 1, 2, \ldots, 248$, are given. These points are also shown in a figure.

The uniformity of the distribution of the points p_n has been examined by comparing the variance of the observed numbers of points p_n in equal disjoint squares, into which the unit square has been divided, to their mean. It has also been examined with the aid of a dendritic method so that the fractions of vertices of particular orders have been compared with the corresponding fractions for random points.

The note terminates with some indications as to the practical use of the points p_n.

(S. Zubrzycki)

BOX, G. E. P. & HUNTER, J. S. (Statist. Tech. Research Group, Princeton Univ.) **9.3** (11.4)
Condensed calculations for evolutionary operation programmes—*In English*
Technometrics (1959) **1**, 77-95 (6 references, 1 table, 14 figures)

This paper presents a detailed description of a simple method for calculating estimates of the " effects " and the error standard deviation for a 2^n factorial design with centre points.

Since an evolutionary operation programme is normally run by plant personnel and the calculations done by them, it is desirable to have a straight-forward method which will tend to reduce arithmetical errors.

Examples of an information board and work sheets are given along with detailed numerical examples of their use for 2^2 and 2^3 designs. The 2^3 design is blocked and the calculation procedure is such as to remove the block effect. The range is used as an estimate of the standard deviation and table of correction factors is included.

There is a discussion of the rationale of the computation method presented.

(P. J. Cislak)

2/721

BRADLEY, R. A., WALPOLE, R. E. & KRAMER, C. Y. (Virginia Poly. Inst., Blacksburg, Va.) **9.1** (9.2)
Intra- and inter-block analysis for factorials in incomplete block designs—*In English*
Biometrics (1960) **16**, 566-581 (12 references)

The authors' main objective in this paper is to extend the recent work of Kramer & Bradley [*Ann. Math. Statist.* (1957) **28**, 349-361 and *Biometrics* (1957) **13**, 197-224] to the use of factorials in the several suitable classes of two-associate class, partially balanced incomplete block designs. Their secondary objective is to complete the earlier work through utilisation of inter-block information in combined intra- and inter-block estimators. They consider specifically the recovery of inter-block information in group divisible designs, both intra- and combined intra- and inter-block analyses for latin-square, sub-type L_2 designs and latin-square, sub-type L_3 designs and then discuss other partially balanced incomplete block designs with two associate classes. In each situation tests of significance and estimators of factorial effects are obtained along with efficiencies of contrasts.

Only basic results are given in this paper. The authors work essentially from treatment estimators for the basic designs and show how they yield the necessary information for the analyses for factorials. Usual concepts of experimental design and factorials apply, and it is emphasised that attention in application must be given to choice and blocking of experimental units, applicability of models, estimation of weights in the recovery of inter-block information, and presentation of results including tables and figures supplementary to basic analysis of variance tables. Methods of analysis follow usual patterns and numerical examples are not included.

The material in this paper greatly increases the numbers of incomplete block designs available for use with factorial treatment combinations. Results for balanced incomplete block designs are well known and may also be obtained by setting $\lambda_1 = \lambda_2$ in formulas presented here for group divisible designs.

(R. E. Walpole)

2/722

BROOKS, S. H. & MICKEY, M. R. (CEIR, Inc., Los Angeles, Cal.) **9.3** (4.7)
Optimum estimation of gradient direction in steepest ascent experiments—*In English*
Biometrics (1961) **17**, 48-56 (4 references, 2 tables)

Suppose R is a response surface functionally dependent on n concomitant variables in some region Φ. Let ϕ be any subregion of Φ such that in ϕ, R is remote from its maximum. The method of steepest ascent, proposed by Box & Wilson [*J. R. Statist. Soc.* B (1951) **13**, 1-45] involves the step-wise progression from an initial set of trials in ϕ to a region ϕ_0 of high response. The path of ascent is determined by successive inferences as to the maximum gradient direction based on sets of trials conducted sequentially en route.

It is desirable to assign t trials per set in such a way as to maximise the gain towards ϕ_0 per unit effort. In the absence of experimental error, $t = n+1$ trials per set constitutes the optimal allocation: however, if experimental error is present, then the estimate of gradient direction becomes uncertain and the problem of optimum allocation of trials reappears.

The authors present a proof which indicates that the maximum gain toward ϕ_0 per unit effort is achieved when $t = n+1$ trials per set; regardless of the magnitude of either the experimental variation or the gradient in ϕ. Proof stems from a derivation of $\mathscr{E}(\cos\theta)$ where θ is the angle of discrepancy between the true gradient and an estimated gradient resulting from a set of t trials. Defining $(1/t)\,\mathscr{E}(\cos\theta)$ to be the expected improvement in location per unit effort, then $(1/t)\,\mathscr{E}(\cos\theta)$ is shown to be maximised for $t = n+1$.

Tabular values are given both for $\mathscr{E}(\cos\theta)$ and for the decrease in efficiency associated with $t>n+1$ trials per set.

(J. E. Dunn)

2/723

CORSTEN, L. C. A. (Inst. Res. on Field Crops, Wageningen, Netherlands) **9.1** (0.6)
Proper spaces related to triangular partially balanced incomplete block designs—*In English*
Ann. Math. Statist. (1960) **31**, 498-501 (3 references)

The incidence matrix for an incomplete block design, N, is the matrix with v rows, v is the number of treatments, and b columns, b being the number of blocks, where the typical element n_{ij}, is unity if the ith treatment occurs in the jth block, and is otherwise zero. The non-negative symmetric matrix $Q = NN'$ of order v has elements q_{ij}, where q_{ii} is equal to the number of replicates of the ith treatment, and q_{ij}, $i \neq j$, is equal to the number of blocks in which the ith and the jth treatment occur together.

Knowledge of proper values and spaces of Q is of interest in finding necessary and sufficient conditions for the existence of designs with given sets of parameters. Knowledge of the proper values of Q for several cases, including the triangular designs as given by Connor & Clatworthy [*Ann. Math. Statist.* (1954) **25**, 100-112], has been utilised in the derivation of necessary conditions for the parameters of such designs.

In this note the author explicitly exhibits the proper spaces of Q for triangular designs. Such information led Ogawa [*Ann. Math. Statist.* (1959) **30**, 1063-1071: abstracted in this journal No. **1**/701, **9.1**] to derive other conditions for the existence of such designs.

(B. S. Pasternack)

2/724

DANIEL, C. (116 Pinehurst Ave., New York 33, New York) **9.2** (4.8)
Parallel fractional replicates—*In English*
Technometrics (1960) **2**, 263-268 (2 tables)

This paper discusses 2^{p-q} fractional replicates designed to study the effects of factors upon two response variables, y_1 and y_2, for certain situations in which y_1 and y_2 are affected by different subsets of the factors. Suppose for example that in a 2^4 experiment factor A affects y_1 only, factors B and C affect both y_1 and y_2, and factor D affects y_2 only. This situation is described by the author as " influence pattern " $A-BC-D$ or $1-2-1$. In general, the influence pattern $p_1-p_2-p_3$ indicates that p_1 factors affect y_1 only, p_2 factors affect both y_1 and y_2, and p_3 factors affect y_3 only. For p_1-p_3, the influence pattern is called a " symmetric influence pattern ".

Some 8-run and 16-run fractional replicate plans are discussed for a few symmetric influence patterns. Two tables of symmetric pattern plans are also presented, one table for 8-run plans and the other for 16-run plans. For each influence pattern tabled the following are given:

(i) the generators of the alias subgroups, plus other important subgroup members in parentheses,
(ii) the generators of the principal blocks.

All of the plans tabled have the following properties:

(1) All confoundings of main effects with two-factor interactions are zero.
(2) For 16-run plans, two-factor interactions are not confounded with each other.
(3) For the 8-run plans one of the following conditions is satisfied:
   (*a*) two-factor interactions are not confounded.
   (*b*) two-factor interactions are confounded in pairs.

(Helen Bozivich)

2/725

DRAPER, N. R. (Math. Res. Centre, U.S. Army, Madison, Wisconsin) **9.3** (9.2)
Third order rotatable designs in three dimensions—*In English*
Ann. Math. Statist. (1960) **31**, 865-874 (5 references, 3 tables)

In this paper the author discusses the construction of third-order rotatable designs in three dimensions, that is to say, designs involving three factors. He presents the conditions for the existence of third order designs and the two conditions which are necessary for non-singularity. It is shown that a third-order rotatable design is singular if and only if all of its points, excluding centre points, lie on a k dimensional sphere centred at the origin.

The four previously known third-order designs are listed in their original geometric framework and are also translated into the mathematical structure used by the author in constructing several infinite classes of third-order designs. Sequential designs, that is to say, those formed from two second-order rotatable designs in three dimensions, are the author's primary consideration. An example for one infinite class is considered in detail. It is shown that two of the previously known designs are members of this class and that their design parameters are located at the extremes possible for this class.

The author states that he intends to present, in a future report, some specific single designs which are to be selected from the infinite classes developed in this paper and will be presented in a form in which they can be used conveniently by experimenters. See also abstract No. **2**/727, **9.3**.

(J. W. Wilkinson)

2/726

DRAPER, N. R. (Math. Res. Centre, U.S. Army, Madison, Wisconsin) **9.3** (9.2)
A third order rotatable design in four dimensions—*In English*
Ann. Math. Statist. (1960) **31**, 875-877 (2 references, 1 table)

In this paper the author presents a third-order rotatable design in four dimensions, that is to say, a design for four factors, which requires 96 points plus n_0 centre points where n_0 may be zero.

This design is a combination of four second-order rotatable arrangements, each containing 24 points. This permits the use of four convenient blocks of equal size or, alternatively, sequential performance in several ways. The notation and definitions of a paper by the author in the same issue are used in obtaining the design; see abstract No. **2**/726, **9.3**.

It is mentioned that previously only one third-order rotatable design in four dimensions was known and that design required 128 points plus centre points.

(J. W. Wilkinson)

2/727

GRAYBILL, F. A. & SESHADRI, V. (Oklahoma State University, Stillwater) **9.1** (7.3)
On the unbiasedness of Yates' method of estimation using interblock information—*In English*
Ann. Math. Statist. (1960) **31**, 786-787 (2 references)

This refers to the results in a previous article [*Ann. Math. Statist.* (1959) **30**, 799-805: abstracted in this journal No. **1**/474, **9.1**]. A proof is given of the unbiasedness of Yates' procedure of using the estimator based on recovery of interblock information if the estimate of the block variance component is positive, and the randomised complete blocks estimator if this estimate is negative.

(R. L. Anderson)

2/728

HOFFMAN, A. J. (G.E.C., Richland, Washington) **9.1** (0.6)

On the uniqueness of the triangular association scheme—*In English*

Ann. Math. Statist. (1960) **31**, 492-497 (6 references)

It has been shown by Connor [*Ann. Math. Statist.* (1958) **29**, 262-266] that the relations that exist among the parameters of the triangular association scheme themselves imply the scheme if $n \geqq 9$.

Shrikhande [*Ann. Math. Statist.* (1959) **30**, 39-47: abstracted in this journal No. **1**/292, **9.1**] extended this argument by demonstrating that this result is also true for $n \leqq 6$: the problem has no meaning for $n < 4$.

The aim of the author in this paper is to show that the result is valid if $n = 7$, but that it is invalid if $n = 8$. For the case $n = 8$, the author proves by the construction of a counter-example that the relations among the parameters of the triangular association scheme do not necessarily imply the triangular association scheme.

The results of this paper have also been obtained through the use of different methods by Chang [*Science Record* (1959) New Series, **3**, 604-613]. Chang has also shown there are exactly three counter-examples when $n = 8$, see [*Science Record* (1960) New Series **4**, 12-18].

(B. S. Pasternack)

2/729

KISHEN, K. (Department of Agriculture, Lucknow) **9.2** (9.1)

A note on the construction of the $(2^{16}, 2^{11})$ and other associated confounded designs keeping up to second interactions confounded—*In English*

J. Ind. Soc. Agric. Statist. (1959) **11**, 180-185 (3 references, 2 tables)

Let $m_t(r, s)$ be the maximum value of m for which it is possible to have a symmetrical factorial design (s^m, s^r); that is, a design with m factors each at s levels, accommodated in blocks of size s^r, so that degrees of freedom belonging to a main effect or interaction involving t or less factors are not confounded. Bose [*Sankyā* (1947) **8**, 107-166] and Rao [*J. R. Statist. Soc.* B (1947) **9**, 128-140] showed that $m_3(r, 2) = 2^{r-1}$ and from this it follows that in a factorial experiment in which each factor is at two levels, and block size is 2^5, the maximum number of factors that can be accommodated so that no main effect or first order interaction is confounded is 16.

In this note the author, using Bose's method, constructs the $(2^{16}, 2^{11})$ design, by selecting 16 points in $PG(4, 2)$, no three of which are collinear. The intra block sub-group of the 32 combinations in the key block is given in a table.

From the $(2^{16}, 2^{11})$ design, the author gets $(2^{15}, 2^{10})$ $(2^{14}, 2^9)$, ... $(2^7, 2^2)$, $(2^6, 2)$ designs by omitting columns successively. The numbers of degrees of freedom confounded for the various higher order interactions in these designs are given in a further table.

(C. S. Ramakrishnan)

2/730

MIKHAIL, W. F. (University of North Carolina, Chapel Hill) **9.1** (-.-)
An inequality for balanced incomplete block designs—*In English*
Ann. Math. Statist. (1960) **31**, 520-522 (2 references)

In this paper the author discusses a balanced incomplete blocks design with v treatments in b blocks of size $k<v$, each treatment appearing in r blocks. It is proven that if $v = nk$, n an integer, $b \geqq v+r-1$.

This is a weaker condition than the usual one that the b blocks can be divided into r sets of n blocks each, so that in each set every treatment occurs exactly once.

(R. L. Anderson)

2/731

PAVATE, M. V. (Central Tobacco Res. Inst., Rajahmundry, India) **9.1** (7.1)
Combined analysis of balanced incomplete block designs with some common treatments—*In English*
Biometrics (1961) **17**, 111-119 (3 references, 3 tables)

The purpose of this paper is to suggest a simplified method to obtain adjusted treatment components for the combined analysis when the individual experiments are laid out in balanced incomplete block designs. Gomes & Guimares [*Biometrics* (1958) **14**, 521-526: abstracted in this journal No. **1**/113, **9.7**] have considered the case when the individual experiments are laid out in randomised complete block designs.

The author presents briefly the notation used. C denotes the number of treatments common to all experiments. The adjusted treatment sum-of-squares for the combined analysis is found by two methods. The general method suggested by Rao [*J. Amer. Statist. Ass.* (1947) **42**, 541-561] is outlined. This method treats the combined data as that of a new incomplete block experiment. Hence the problem of combined analysis reduces to the solution of a set of adjusted normal equations.

The simplified method, which the author introduces, makes use of the individual analysis already performed on the balanced incomplete block experiments. This method is presented in some detail. Estimation of treatment estimates and variances of treatment differences is also discussed. Four special cases of combined analysis of these balanced incomplete block experiments are discussed.

(*a*) Experiments with only one common treatment which might be a control. In this case $C = 1$.
(*b*) Several experiments with identical parameter values and treatments. In this case $C = v$ where v is the number of treatments in an individual block.
(*c*) & (*d*) Cases where balanced incomplete block designs reduce to randomised complete block designs with C common treatments. These cases are considered in detail by Gomes & Guimares.

Finally the author presents an example dealing with two experiments: the individual analysis of both is discussed. The combined analysis is by the author's simplified method.

(J. J. Bartko)

2/732

RAGHAVARAO, D. (University of Bombay) **9.1** (0.4)

A generalisation of group divisible designs—*In English*

Ann. Math. Statist. (1960) **31**, 756-771 (9 references, 10 design plans)

Group divisible designs were first introduced by Bose & Connor [*Ann. Math. Statist.* (1952) **23**, 367-383], and were later extended to *m*-associate classes by P. M. Roy [*Science and Culture* (1953) **19**, 210-211]. Roy called these designs " Hierarchial Group Divisible designs with *m*-associate classes ". Further work in this direction was not pursued in the literature until the publication of this paper in which the author studies these designs systematically.

In the second section the author gives a compact definition of a group divisible *m*-associate design in which the parameters of the design are presented in a slightly different form than that of Roy. In Section three the author demonstrates the uniqueness of the association scheme from the parameters of the design. The designs are divided into $(m+1)$ classes in Section four, and in Section five some interesting combinatorial properties of these designs are discussed. The following section offers a presentation of some known results concerning the Legendre symbol, the Hilbert norm residue symbol and the Hasse-Minkowski invariant. The final section of this paper describes necessary conditions for the existence of symmetrical regular group divisible *m*-associate designs, and numerical illustrations of these designs are given in the Appendix.

(B. S. Pasternack)

2/733

RAGHAVARAO, D. (University of Bombay) **9.8** (9.1)

Some aspects of weighing designs—*In English*

Ann. Math. Statist. (1960) **31**, 878-884 (9 references)

The author supposes we have N objects to be weighed in N weighings with a balance having no bias. Let the design matrix $X(N\times N) = (x_{ij})$ where $x_{ij} = 1\,(-1)$ if the jth object is placed in the left or alternatively the right pan in the ith weighing, and $x_{ij} = 0$ if the jth object is not weighed in the ith weighing. In a previous paper by the author " Some optimum weighing designs " [*Ann. Math. Statist.* (1959) **30**, 295-303: abstracted in this journal No. **1**/288, **9.8**] it was shown that the design matrices $X = P_N$ and $X = S_N$ are the most efficient weighing designs under Kishen's definition of efficiency, see " On the design of experiments for weighing and making other types of measurements " [*Ann. Math. Statist.* (1945) **16**, 294-300]. This is related to the magnitude of the trace of the matrix $(X'X)^{-1}$, when N is odd and $N \equiv 2 \pmod 4$, respectively, subject to the conditions that the variances of the estimated weights are equal and the estimated weights are equally correlated. Here $P_N'P_N = (N-1)I_N + E_{NN}$ and $S_N'S_N = (N-1)\,I_N$ where I_N is the identity matrix of order N and E_{NN} is an Nth order matrix with positive unit elements everywhere.

In this paper it is shown, subject to the same conditions as stated above, that the P_N matrices are the best weighing designs under the definitions of Mood, see " On Hotelling's weighing problem " [*Ann. Math. Statist.* (1946) **17**, 432-446] and Ehrenfeld, " On the efficiencies of experimental design " [*Ann. Math. Statist.* (1955) **26**, 247-255], when N is odd; while the S_N matrices are the best weighing designs under the definition of Ehrenfeld when $N \equiv 2 \pmod 4$. Under Mood's definition of efficiency, the best weighing design X, when $N \equiv 2 \pmod 4$, is shown to be that for which $X'X = (N-2)I_N + 2E_{NN}$. The best weighing design X according to Mood's definition is the one for which the determinant of $X'X$ is maximum. The best weighing design according to Ehrenfeld's definition is the one for which the smallest characteristic root of $X'X$ is maximum.

It is shown that the necessary condition for the existence of a weighing design X such that $X'X = (N-2)\,I_N + 2E_{NN}$ is that $N = \{4+(3f^2+4)^{\frac{1}{2}}\}/3$, where f is an integer. For $N<200$, this implies the existence of such weighing designs only for $N = 6$ and $N = 66$. The author also shows the impossibility of S_N matrices of order $N = 22$, 34, 58 and 78.

(J. W. Wilkinson)

2/734

RAO, P. V. (College of Science, Nagpur, India) **9.1** (0.6)
The dual of a balanced incomplete block design—*In English*
Ann. Math. Statist. (1960) **31**, 779-785 (9 references)

Shrikhande [*Biometrics* (1952) **8**, 66-71] and P. M. Roy [*Sankhyā* (1954) **14**, 39-52] have shown that certain balanced incomplete block designs can be dualised to give partially balanced incomplete block designs with exactly two associate classes. J. Roy & Laha [*Sankhyā* (1956) **17**, 115-132] have obtained a necessary and sufficient condition for the dual of a balanced incomplete block design to be a partially balanced incomplete block design with two associate classes.

In this paper, a general result regarding the dual of a balanced incomplete block design is established, and the results of Shrikhande and P. M. Roy are obtained as particular cases. An illustration to show the use of the result when the dual is not a 2-associate partially balanced incomplete block design is also given.

(P. V. Rao)

2/735

SCHWARZ, G. (Columbia University, New York) **9.2** (7.4)
A class of factorial designs with unequal cell-frequencies—*In English*
Ann. Math. Statist. (1960) **31**, 749-755 (8 references)

A class of multifactorial designs are defined and analysed. The designs considered each have a total number of observations that cannot be divided equally among the cells of the designs; however, by distributing the observations in a manner that is in a certain sense symmetrical, the equations that determine the least-squares estimates of the linear parameters become explicitly solvable.

The first case which the author treats in this paper is the case of two non-interacting factors with arbitrary numbers of levels. In the *n*-factor case, he restricts himself to the discussion of factors having an equal number of levels. After defining the designs, the estimates are computed. Some general discussions of the symmetries and algebraic properties involved conclude the paper.

(G. Schwarz)

2/736

SHAH, B. V. (University of Bombay) **9.1** (0.6)

A matrix substitution method of constructing partially balanced designs—*In English*

Ann. Math. Statist. (1960) **31**, 34-42 (6 references)

Given an incomplete blocks design for v treatments and b blocks with the following incidence matrix, $N = (n_{ij})$, where n_{ij} is the number of times the ith treatment occurs in the jth block. The author is concerned only with designs for which $n_{ij} = 0$ or 1 and

$$\sum_{j=1}^{b} n_{ij} = r < b \quad \text{and} \quad \sum_{i=1}^{v} n_{ij} = k < v.$$

If each treatment appears with every other treatment in λ blocks, that is to say $\sum_j n_{ij} n_{lj} = \lambda$ for $l \neq i$, the design is balanced; if certain treatments appear together in λ_1 blocks, others together in λ_2 blocks, ..., others in λ_m blocks, the design is partially balanced with m associate classes.

Vartak [*Ann. Math. Statist.* (1955) **26**, 420-438] considered the construction of experimental designs utilising kronecker products of incidence matrices. His method essentially consisted of the replacement of two elements, 0 and 1, by two matrices. Shah, in an earlier paper [*Ann. Math. Statist.* (1959) **30**, 48-54: abstracted in this journal No. **1**/291, **9.1**] generalised this approach by using the incidence matrices of balanced incomplete block designs for substitution. In the present paper the author continues his generalisation by considering the incidence matrices of partially balanced incomplete block designs and factorial experimental designs as replacement elements.

The author introduces fundamental notions concerning canonical matrices with reference to the partially balanced incomplete block designs in Sections 2 and 3 followed by a discussion of associable designs and their properties in Section 4. In Section 5 balanced matrices are defined and in Section 6 a method is presented for the construction of designs by substitution of the incidence matrices of a balanced matrix. Finally, the application of this method to the construction of factorial experimental designs is given in Section 7.

The author indicates that these methods may be of considerable importance in the construction of confounded asymmetrical factorial designs when there are many factors.

(B. S. Pasternack)

2/737

SHAH, B. V. (University of Bombay) **9.2** (9.1)

Balanced factorial experiments—*In English*

Ann. Math. Statist. (1960) **31**, 502-514 (15 references)

Factorial experiments are usually of such a nature as to prohibit all possible treatment combinations from appearing in a block: it is necessary, therefore, to consider methods of confounding. The author in an earlier paper [*Ann. Math. Statist.* (1958) **29**, 766-779] considered the problem of achieving " complete balance " over various interactions in factorial experiments. " Complete balance " is achieved over an interaction, if and only if all the normalised contrasts belonging to the same interaction are estimated with the same variance. In the present paper, the author considers a class of factorial experiments which are defined as balanced factorial experiments. An experiment is called a balanced factorial experiment, if the following conditions are satisfied:

(i) Each of the treatments is replicated the same number of times.
(ii) Each of the blocks has the same number of plots.
(iii) Estimates of contrasts belonging to different interactions are uncorrelated with each other.
(iv) " Complete balance " is achieved over each of the interactions.

In the fifth section of this paper Shah proves several theorems which indicate a method of analysis for balanced factorial experiments which is particularly useful when estimates of interactions at different levels are required. A method for constructing these experiments is given in the final section.

(B. S. Pasternack)

2/738

VARTAK, M. N. (University of Bombay) **9.1** (0.6)

Relations among the blocks of the kronecker product of designs—*In English*

Ann. Math. Statist. (1960) **31**, 772-778 (7 references)

For certain incomplete block designs interesting relations among blocks have been found. Fisher [*Ann. Eugenics* (1940) **10**, 52-75] has shown that in the case of a symmetrical balanced incomplete block design with parameters $v = b$, $r = k$, λ, any two blocks have exactly λ treatments in common. For an affine resolvable balanced incomplete block design with parameters

$$v = nk = n^2\{(n-1)t+1\},\quad b = nr = n\{n^2t+n+1\},$$

$\lambda = nt+1$, Bose [*Sankhyā* (1942) **6**, 105-110] has shown the blocks can be divided into sets of n blocks, such that each set is a complete replication and any two blocks have either $k^2/v = (nt-t+1)$ or nil treatments in common according to whether they belong to different groups or the same group. Similar results have been obtained by Connor [*Ann. Math. Statist.* (1952) **23**, 57-71] and Bose & Connor [*Ann. Math. Statist.* (1952) **23**, 367-383].

In this paper the author confines his attention to partially balanced incomplete block designs having two or three associate classes with the objective of obtaining analogous relationships among blocks of balanced incomplete block designs, as those described above, for the blocks of their kronecker product. The interconnection of the results he obtains and the theorems on inversion of designs, P. M. Roy [*Bull. Cal. Math. Soc.* (1954) **46**, 47-58] are also discussed.

In one main theorem, the author proves that if for each of two designs with incidence matrices N_1 and N_2, there exists a pair of blocks with, respectively, M_1 and M_2 treatments in common, the design with incidence matrix based on the kronecker product $N_1 \times N_2$ will have a pair of blocks with M_1M_2 treatments in common.

Other theorems and corollaries show how to classify the method of the classification of blocks according to the number of treatments that they have in common with some other block.

(B. S. Pasternack)

2/739

BALOGH, T. (Kossuth University, Debrecen) **10.8** (0.6)
Generalisation of a theorem of U. Grenander—*In English*
Second Hungarian Mathematical Congress (1961) (4 references)

The extension of the theory of stochastic processes to the case of matrix-valued variables made it necessary to determine the minimum of the spur of the so-called generalised Toeplitz-form

$$\frac{1}{2\pi}\int_0^{2\pi} \mathbf{P}_n(z) f(x) \mathbf{P}_n^*(z)\, dx; \qquad z = e^{ix}$$

under the condition $\mathbf{P}_n(\alpha) = E$, where $\mathbf{P}_n(z)$ is a trigonometric polynomial matrix, and $f(x)$ is a positive definite Hermitian functional matrix.

In the case of ordinary Toeplitz-forms the above problem has been solved by Szegö. The method by which we obtain our result is essentially a generalisation to the case of matrices of the proof of Szegö. If we apply the theorem so obtained to matrix-valued stochastic processes, we obtain a natural generalisation of the theorem of Grenander.

(T. Balogh)

2/741

BARLOW, R. E. & HUNTER, L. C. (Sylvania Electric Defense Lab.,
Mountain View, Cal.) **10.1** (10.2)
System efficiency and reliability—*In English*
Technometrics (1960) **2**, 43-54 (7 references, 5 figures)

This paper presents a model useful in determining the reliability and efficiency of large, complex systems. Repair is an integral part of the model and the assumption of component independence need not be made.

A system operating in time can be in any one of many states. At a given time the system will be in one of the states. Thus the state of the system is essentially a random variable and as a function of time it is a stochastic process $[X(t),\ 0 \leqq t < \infty]$. Usually it is advantageous for the system to be in a certain class of states. The designation of these states is made by the definition of a gain function g defined on the state-space of the system. Since the state-space of the system as a function of a time is a stochastic process, the gain function is composed with the process; that is, $g\,[X(t)]$. The reliability of the system at time t, $R(t)$, is defined as the expected value of the gain function. For example, if $g\,[X(t)] = 1$ when the system is in a favourable state and $g\,[X(t)] = 0$ otherwise, the reliability is the probability that the system is in a favourable state at time t.

The final factor in the model deals with the possibility that the system is more likely to fail at one time than another. Thus the environment may provide a distribution function $F(t)$ on the time axis. Efficiency is defined as the expected value of the reliability function. If the system is operating under constant threat, efficiency is the time average of the reliability function. The limiting efficiency is defined in terms of the time average.

The model presented is very general and permits the use of probabilistic techniques developed for stochastic processes. However, actual computations of system efficiency and reliability are easily accomplished only for systems with finite state spaces and for those which can be described by Markoff processes.

(A. P. Berens)

2/742

BHAT, B. R. (University of Western Australia, Nedlands) **10.2** (4.1)

Maximum likelihood equation for positively regular Markoff chains—*In English*

Sankhyā (1960) **22**, 339-344 (8 references)

In this paper the author considers the problem of maximum likelihood estimation of an unknown parameter entering in the specification of the transition probability matrix of a simple stationary and positively regular Markoff chain. The main results can be stated as follows: let $\pi(\theta)$ denote the transition probability matrix when the true value of the parameter is θ and p, the estimate of π based on the empirical distribution of n successive observations. Let T be any Fisher-consistent estimate of θ based on p. The variance of T then satisfies the well-known Cramér-Rao information inequality under certain regularity conditions. If θ^* denotes an estimator of θ which makes the likelihood a maximum and

$$\inf_{|\theta-\theta_0|>\delta} \left[-\sum_j \pi_{ij}(\theta_0)\{\log \pi_{ij}(\theta)/\pi_{ij}(\theta_0)\}\right] = \varepsilon > 0$$

for every $\delta > 0$ and all i, θ_0 being the true value of the parameter, then θ^* is a consistent estimate of θ. Under certain conditions (which are not stated in the paper) the uniqueness of the root of the likelihood equation, at which the maximum is attained, and the asymptotic normality of its distribution are asserted. All these results are generalisations of those of Rao [*Sankhyā* (1957) **18**, 139-148] in the case of independent observations.

(K. R. Parthasarathy)

2/743

BLOM, G. (University of Lund, Sweden) **10.1** (-.-)

Hierarchical birth and death processes. I. Theory—*In English*

Biometrika (1960) **47**, 235-244 (9 references)

A Markoff chain is defined with certain transition rates, $m-1$ indices for " births " and m indices for " deaths ". Thus transitions occur only between states with one more or less indices, and only one index at a time is either deleted or added. Each index is restricted to the same finite set of values 1 to N.

Steady state probabilities are found for the finite $(m \leq n)$ " last come, first served " process in which only the index at the right-hand end can change. It is shown that the same solutions are obtained in the " first come, first served " process in which additions to the indices are at the right-hand end and deletions at the left-hand end, provided that $q_1 = q_2 = \dots = q_N$.

A sufficient set of conditions is given for the general infinite processes to tend to a steady state. The steady state probabilities are shown to be the same as those derived earlier, for Type A processes in which

$$q_1 = q_2 = \dots - = q_N$$

and for Type B processes in which $a_{m-1,\nu} = c_m b_{m,\nu}$ where $c_m > 0$ and $\nu = 1, \dots, m$.

In a second paper the author discusses some practical applications of this theory. See abstract No. **2**/745. **10.1**.

(A. S. Miller)

2/744

BLOM, G. (University of Lund, Sweden) **10.1** (-.-)

Hierarchical birth and death processes. II. Applications—*In English*

Biometrika (1960) **47**, 245-251 (5 references)

The theory of hierarchical birth-and-death processes is discussed in Part I of the paper. See abstract No. **2**/744, **10.1**.

It is shown in this second part of the paper that this theory may be applied to the machine interference problem with different rates of breakdown and different service times for different machines and with either one or several operators.

The theory is also applied to a telephone traffic problem. A group of n channels is used by N subscribers. Subscribers make calls at random but at different rates and the durations of their calls have (i) exponential distributions with varying parameters, (ii) mixed exponential distributions. If all n channels are occupied no queueing is allowed, the subscriber makes another attempt later.

(A. S. Miller)

2/745

CIUCU, G. (University " C. I. Parhon ", Bucharest) **10.1** (1.5)

I. Some properties of chains with complete connections

II. The law of large numbers for chains with complete connections—*In Rumanian*

Com. Acad. R.P.R. (1960) **10**, 561-563, 643-645 (7 references)

In these papers concerning chains with complete connections the author considers a chain with a denumerable set of states Ω, defined by the transition probabilities

$$P_{1,l}(c;\ A^{(l)}),\quad c\in \prod_{j\in -N}\Omega^{j},\qquad \Omega^{j}=\Omega\,(j\in -N),$$

$$-N=\{\ldots,\ -1,0\},\quad A^{(l)}\subset\Omega^{(l)}=\prod_{j=1}^{l}\Omega^{j},$$

$$\Omega^{j}=\Omega\,(1\leqq j\leqq l).$$

A chain of this kind is of type (G), if there exists a function $P^{\infty}_{1,l}(A^{(l)})$, defined for all $A^{(l)}\subset\Omega^{(l)}$ ($l\in N=\{1, 2, \ldots\}$) and two constants L, $\lambda>0$ such that

$$\left| P^{(n)}_{1,l}(c;\ A^{(l)})-P^{\infty}_{1,l}(A^{(l)})\right|\leqq L\exp(-\lambda\sqrt{n}),$$

where $P^{(n)}_{1,l}(c;\ A^{(l)})$ is the n-step transition probability.

In the first note it is shown that for chains with complete connections of type (G) a central limit theorem and the theorem of the iterated logarithm hold true; in the second note the law of large numbers is considered.

(R. Theodorescu)

2/746

COX, D. R. (Birkbeck College, London) **10.1** (2.7)

On the number of renewals in a random interval—*In English*

Biometrika (1960) **47**, 449-452 (7 references)

A renewal process is defined by the author in this paper as a sequence $(X_1, X_2, \ldots)$ of independent identically distributed random variables, taking only positive values, and is assumed here to have an absolutely continuous distribution. If X_i is interpreted as the lifetime of the ith component to be used, renewals occur at times $X_1, X_1+X_2, \ldots$. The random variable N_t is defined as the number of renewals $(0, t)$.

There is an extensive literature both on the exact properties of N_t and on asymptotic results which hold as $t \to \infty$. The exact results involve the inversion of a Laplace transform, so that simple explicit results are obtained only for especially simple distributions of the X_i.

Suppose now that T is a positive random variable distributed independently of the X_i: let N be the number of renewals in $(0, T)$. It is pointed out that when T has a Γ distribution, the properties of N can be obtained without the inversion of a Laplace transform. Hence, particularly when the index of the Γ distribution is a small integer, simple explicit results can be obtained. A number of special cases are worked out.

(D. R. Cox)

2/747

FINCH, P. D. (London School of Economics) **10.4** (0.1)

Deterministic customer impatience in the queueing system. GI/M/I—*In English*

Biometrika (1960) **47**, 45-52 (5 references)

This paper gives the solution of an integral equation occurring in the analysis of a particular queueing system when the common distribution function $[V(x)]$ of the identically distributed random variables $[V_n]$ takes the form

$$V(x) = \mu e^{\mu x} \int_x^\infty A(y)\, e^{-\mu y}\, dy$$

where $x > 0$. The properties of the queueing system are as follows:

(i) Customers arrive at time instants (t_n) after $t_0 = 0$ so that the $\tau_n = t_n - t_{n-1}$, $(n = 1, 2, \ldots)$ are independent and identically distributed non-negative random variables with common distribution function $A(x)$ and finite expectation

$$\alpha = \int_0^\infty x\, dA(x).$$

(ii) Customers are served in order of arrival and service is immediate if a customer arrives and finds the server free.

(iii) A customer will wait for service only for a time not exceeding W at which time he departs never to return.

(iv) The service times of customers who receive service are independently and identically distributed random variables which are independent of $\{t_n\}$ and with a common distribution function $B(x) = 1 - e^{-\mu x}$ for $x \geqq 0$.

The case where $\{t_n\}$ is a Poisson process was studied by Barrer [*Operat. Res.* (1957) **5**, 650-656] who obtained the limiting distribution of queue-size at an arbitrary instant of time.

The second section of this present paper contains the solution of the integral equation

$$G(x) = \int_{-\infty}^{x} G(x-y)\, dV(y), \quad 0 < x \leqq W$$

and this is illustrated for the case where $\{t_n\}$ is an Erlang (E_k) process with mean k/λ. The following section relates the analysis to the case of $W = \infty$ and the traffic intensity $I = 1/(\mu\alpha) < 1$. Where $I = 1$, a solution does not exist. The case of $I > 1$, $W = \infty$ is the limiting-time distribution of a queue with Poisson input process with parameter μ and service-time distribution $A(x)$. The last two sections of this paper deal with the limiting distribution of queue size at the instants a service commences and the case of the many server queue.

(Florence N. David)

2/748

DANSKIN, J. M. (Rutgers University, New Jersey, USA) **10.5** (0.7)
The theory of reconnaissance—*In English*
ARL Technical Report 60-337 (1960) Aero. Res. Labs., Wright-Patterson Air Force Base.
iv+20 pp. (2 references)

In a previous paper [*WADC Technical Note* 59-409: abstracted in this journal No. **2**/381, **10.5**] with the same title the author considered the one-sided reconnaissance problem, in which the reconnoiterer seeks to maximise the information or minimise the confusion obtained as a result of the expenditure of a given effort; it being assumed that the side being reconnoitered remains passive. This problem was formulated as a problem in information theory. This report concerns the two-sided reconnaissance problem, in which the side being reconnoitered seeks to minimise the information or maximise the confusion obtained by the reconnoiterer, while maintaining at least a certain minimum acceptable threat with a fixed budget. This problem is formulated as a zero-sum, two-person game: for example consider a defender trying to protect ballistic missile installations against reconnaissance. When planning at a period substantially preceding the period in which he will be reconnoitered, he is limited by a budget but not by equipment, and may choose to divert money to, for instance, concealment or decoys. If the period in which the reconnaissance is to take place is at hand, he will be more restricted, for example perhaps to moving a limited number of missiles from site to site in order to make empty sites decoys. Concealment will be fixed, since this involves large capital expense. The latter situation, equipment limitation, is a special case of the former budget limitation. A solution is given for the special but not the general case, though it is proved that there exists a solution in mixed strategies for the latter.

A model is given for evaluating the effectiveness of a human photointerpreter, or a photo-interpreting system; this model separates the confusion due to the photointerpreter (human or system) from that due to the photographs. The expected confusion to an interpreter using the system R is denoted by C_R, the confusion inherent in the photographic system used by C_P. Formulas are given for C_R and C_P; it is proved that $C_R \leqq C_P$. The minimum achievable confusion is C_P: the measure of effectiveness of the photo-interpreting system is defined as $E = C_R - C_P$, the additional confusion introduced by that system over that inherent in the photographs.

continued

DANSKIN, J. M. (Rutgers University, New Jersey, USA) **10.5** (0.7)
The theory of reconnaissance—*In English* *continued*
ARL Technical Report 60-337 (1960), Aero. Res. Labs., Wright-Patterson Air Force Base.
iv+20 pp. (2 references)

An error in *WADC Technical Note* 59-409, in the statement (Section 11, page 30) about the solution for the concave-convex case, is pointed out and corrected: the results are not invalidated.

(H. L. Harter)

FISZ, M. (Math. Inst., Polish Acad. Sci., Warsaw) **10.0** (-.-)
Remarks on the sample functions of some stochastic processes—*In English*
Bull. Acad. Polon Sci. III (1960) **8**, 355-358

The author considers the real separable stochastic process $\{x_t, t \in I_0\}$, where I_0 is a closed finite interval. Denote further by x_I the increment of x_t in the interval I and by $|I|$—the length of I. Consider the following functions:

$$a(I) = \Pr(x_I \neq 0), \quad b(I, \varepsilon) = \Pr(x_I| > \varepsilon),$$

$$A(I) = \int_{\mathcal{T}} a(I),\ B(I, \varepsilon) = \int_{\mathcal{T}} b(I, \varepsilon),\ \ Q(t) = \lim a(I)/|I|$$

as I contracts to a fixed point $t \in I$. $A(I)$ and $B(I, \varepsilon)$ are the Burkill integrals of $a(I)$ and $b(I, \varepsilon)$, respectively. The author characterises the sample functions of the process $\{x_t, t \in I_0\}$ under different assumptions concerning above-mentioned functions.

(L. Kubik)

2/751

GHOSAL, A. (Central Fuel Research Institute, Dhanbad, India) **10.4** (10.1)
Emptiness in the finite dam—*In English*
Ann. Math. Statist. (1960) **31**, 803-808 (8 references)

This paper discusses the general problem of emptiness in the finite dam and considers the probability that, starting with an arbitrary storage, the dam dries up before it fills completely. Both discrete and continuous inputs are considered.

In the model considered, the storage Z_t defined for $t = 0, 1, 2, \ldots$, is given by

$$Z_{t+1} = \begin{cases} Z_t + X_t - m & \text{if } m < Z_t + X_t < k \\ 0 & \text{if } Z_t + X_t \leqq m \\ k - m & \text{if } Z_t + X_t \geqq k \end{cases}$$

where the X_t are identically distributed independent inputs, which flow into the dam in the intervals $(t, t+1)$. Defining V_i as the conditional probability that, starting with storage i, the dam becomes empty before it fills completely an exact solution for V_i is obtained under the assumption of a geometric input, $\Pr(X_t = j) = ab^j$; $b = 1 - a$, $j = 0, 1, 2, \ldots$.

The continuous analogue of V_i, $V(y)$ is defined and exact solutions for $V(y)$ are obtained under the assumption of an exponential input, $g(X) = \mu \exp(-\mu X)$, $X > 0$ and a gamma input,

$$g(X) = \frac{\mu^p}{(p-1)!} \exp(-\mu X) X^{p-1}, \quad X > 0.$$

It is shown that if $H(y)$ is the stationary continuous distribution function of the dam content, then $V(y) = H(k-y)$ for the continuous inputs considered. This relation also holds for a discrete input.

(P. L. Meyer)

2/752

HAIGHT, F. A. (University of California, Los Angeles) **10.4** (2.5)
Queueing with balking. II—*In English*
Biometrika (1960) **47**, 285-296 (8 references)

In a previous paper [*Biometrika* (1957) **44**, 360-369] the author considered the queueing problem where the rate of arrivals may be greater than the service rate. However, arrivals join the queue only if its length is K or less otherwise they balk, i.e. go away and do not return. The values of K chosen by successive arrivals are regarded as being random samples from the balking distribution. The inter-arrival time and service time distributions are negative exponential and there is a single server.

In his paper the author considers queue lengths and waiting-time distributions under certain specific conditions and also the probability distribution for the time gap between an arrival (whether joining or not) and the previous joiner. The corresponding quantities for gaps between a joiner and the preceding joiner are also found.

Finally a numerical example is given based on a simulated queue using a Poisson balking rule. The observed numbers of negative exponential phases in inter-arrival and inter-joining gaps are given and geometric, negative-binomial and Borel distributions have been fitted.

(C. Burrows)

2/753

IOSIFESCU, M. & THEODORESCU, R. (Central Statistical Office, and Inst. Math., Bucharest) **10.1** (1.0)
On certain linear chains with complete connections—*In Russian*
Rev. Math. Pures Appl. (1961) **6**, 167-170.

The problem of the ergodic structure of the chains with complete connections was first considered by Fortet in his thesis for the alternative linear case. In the present paper the authors provide adequate conditions under which a linear chain with complete connections with a finite set of states, containing more than two states, can be considered as an alternative linear chain. For details on these processes see the recent monograph by Ciucu & Theodorescu [*Processes with complete connections* (1960) Bucharest: Ed. Acad. R.P.R.].

(G. Sîmboan)

2/754

LAMPERTI, J. (Stanford University) **10.1** (-.-)

The first-passage moments and the invariant measure of a Markoff chain—*In English*

Ann. Math. Statist. (1960) **31**, 515-517 (3 references)

The author considers an irreducible recurrent Markoff chain and develops a simple expression for the $(k+1)$th factorial recurrence-time moment for state 0 in terms of the factorial first-passage time moments for all positive states and the invariant measure of the chain, valid for $k = 0, 1, 2, \ldots$. This generalises the well-known formula for $k = 0$ and the formula for $k = 1$ due to Chung [*Trans. Amer. Math. Soc.* (1954) **76**, 397-419].

(W. J. Hall)

LESLIE, P. H. & GOWER, J. C. (Bur. Animal Pop., Oxford and Rothamsted Exp. Sta.) **10.9** (-.-)

The properties of a stochastic model for the predator-prey type of interaction between two species—*In English*

Biometrika (1960) **47**, 219-234 (9 references, 4 tables, 2 figures)

A model is considered in which the birth rates (per individual, per unit time) of two species, prey and predator, are constant while the death rate of the prey is approximately linear in both the numbers of prey and predator and the death rate of the predator is approximately linear in the ratio of predator to prey. The model is in fact stated, for computational purposes, in terms of the means and variances of the numbers of each species (assumed to be normally distributed before rounding off to the nearest non-negative integers) present at time $t+1$, given the numbers of each species at time t.

Approximate theoretical variances and covariances for the fluctuations of the process in the neighbourhood of the stationary state are given and are shown to compare quite favourably with the results from Monte Carlo realisations of the process. Chances of extinction of either species are also discussed.

A second model is then considered in which only a proportion k of the prey population is exposed to risk. This model is treated similarly to the first.

(A. S. Miller)

LUKASZEWICZ, J. (Wrocław University, Wrocław) **10.4** (10.1)
A remark on a theorem by Khintchine from the theory of random streams—*In Russian*
Coll. Mat. (1960) **7**, 285-287

The author notices that a lemma in Khintchine's book [*Mathematical Methods in the Theory of Queueing* (1960) London: Griffin] stating that for every Palm stream, that is for every stationary orderly stream with a limited after-effect, and for every $r = 1, 2, \ldots,$ we have

$$\lim_{u \to 0} \psi_{r+1}(u)/\psi_r(u) = 0,$$

where $\psi_r(u)$ is the probability of having at least r calls in a time segment of length u, is true only under the additional assumption that $\psi_r(u) > 0$ for $u > 0$.

Moreover, he gives a generalised version of this lemma which is true for all Palm streams without exception and enables us to complete the proof of a theorem about probability distribution of distances between consecutive calls in Palm streams in its full generality; as it is stated in Khintchine's book.

(S. Zubrzycki)

2/757

MEERENDONK, H. W. van der & BOS, G. G. J. (N.V. Philips, Eindhoven) **10.4** (-.-)
Rules for the classification of seasonal production—*In Dutch*
Statist. Neerlandica (1961) **15**, 177-187 (2 tables, 1 graph)

This paper deals with the analysis of a problem often met when determining purchasing policies. From a number of different kinds of goods, which are to be bought, the total quantity is restricted by storing capacity or financial limits. The problem " how much of each kind will be bought " is solved for a discrete as well as for a continuous distribution of demand for the various goods.

(H. W. van der Meerendonk)

2/758

NELSON, E. & VARBERG, D. (University of North Dakota & Hamline University, Minn.) **10.1** (1.2)
Expectations of functionals on a stochastic process—*In English*
Ann. Math. Statist. (1960) **31**, 574-578 (6 references)

Let $x\{t\}, 0 \leqq t < \infty$ be the separable and centred infinitely divisible stochastic process which gives the sum of $N(t)$ independent Bernoulli random variables with equally likely values ± 1; where $N(t)$ is an independent Poisson process. The characteristic function of this process is then

$$\mathscr{E}\{\exp[i\xi x\{t\}]\} = \exp[at(\cos\xi - 1)].$$

Let V be a non-negative real function and set

$$\Psi_n(u) = \int_0^\infty e^{-st}\mathscr{E}\left\{\exp\left[-u\int_0^t V\{x(\tau)\}\,d\tau\right]I(t, n)\right\}dt$$

where $I(t, n)$ equals 1 or 0 according as $x\{t\} = n$ or $\neq n$.

The main theorem of this paper states that Ψ_n satisfies the difference equation

$$\Psi_{n+1} - (2/a)[s + a + uV(n)]\Psi_n + \Psi_{n-1} = -(2/a)\delta_{n,0},$$

the quantity $\delta_{n,0}$ being the kronecker delta. The method and notation follow those of Kac " On some connections between probability theory and differential and integral equations " [*Proc. Second Berkeley Symposium Math. Statist. & Probability* (1951) 189-215].

Two examples are then given, firstly $V(x)$ is zero when $a<x<b$ and is unity otherwise; and secondly $V(x) = x^2$.

(R. Pyke)

2/759

PARTHASARATHY, K. R. (Indian Statistical Institute, Calcutta) **10.6** (10.2)
On the estimation of the spectrum of a stationary stochastic process—*In English*
Ann. Math. Statist. (1960) **31**, 568-573 (7 references)

Let $x_1, x_2, \ldots, x_N$ be N observations of a discrete parameter weakly stationary stochastic process; a second order process with stationary mean and covariance functions. Set $\mathscr{E}\, x_t = 0$ and $\rho_n = \mathscr{E}\, x_t x_{t+n}$. Let

$$\hat{\rho}_n = (x_1 x_{1+|n|} + \ldots + x_{N-|n|} x_N)/(N - |n|)$$

be the sample covariances, and consider the following estimates of the spectral distribution F:

$$\hat{F}_N(\lambda) = 2\pi^{-1} \sum_{k=-R(N)}^{R(N)} a_{k,N}(\hat{\rho}_k/ik)e^{ik\lambda}$$

where the summand for $k = 0$ is interpreted to be $a_{0,N}(\lambda + \pi)$ and the $a_{k,N}$ are constants chosen to satisfy firstly, $a_{k,N} \to 1$ as $N \to \infty$, secondly $a_{k,N} = a_{-k,N}$ and thirdly, $\hat{F}_N$ is a distribution function.

Investigations into the asymptotic behaviour of such " weighted periodogram " estimates of the spectral density function have been carried out under various assumptions by Grenander & Rosenblatt [*Statistical Analyses of Stationary Time Series* (1957) New York: Wiley] and by Parzen, [see, for example, " On asymptotically efficient consistent estimates of the spectral density function of stationary time series," *J. R. Statist. Soc.* B (1958) **20**, 303-322: abstracted in this journal No. **1**/138, **10.6**].

" In this paper, the problem of estimating the spectral distribution as well as the spectral density, if it exists, of a weakly stationary process is solved under the sole assumption that the sample covariances converge almost surely and in mean to the true covariances. . . . The existence of estimates which converge uniformly strongly to the spectral density of the process is proved under the assumption that the density has an absolutely convergent Fourier series," states the author. He applies a theorem of Bochner to obtain a uniform bound on the bias of these estimates in the case of a continuous spectral distribution.

(R. Pyke)

2/760

RAO, M. M. (Carnegie Institute of Technology, Pittsburg) **10.2** (-.-)
Estimation by periodogram—*In English*
Trab. Estadíst. (1960) **11**, 123-137 (7 references)

This paper is devoted to establishing the relationship between the methods of estimation of the period and amplitude of a harmonic model by periodogram and maximum-likelihood or least-squares.

Let T be a set of integers and x_t, $t \in T$, be a process which satisfies the relation $x_t = a \cos \omega t + b \sin \omega t + \varepsilon_t$, where a, b and ω are real constants to be estimated from the observations on x_t, and ε_i are " errors ". Assume:

(i) $-\rho < a, \quad b < \rho, \quad 0 < \omega < \rho$.

(ii) The random variables ε_t, $t \in T$, have zero mean and are uncorrelated with a finite positive variance independent of t.

(iii) $T \subset I$, where I is a finite fixed interval.

Let X_t ($t = 1, \ldots, n$) be the observations on the process and define $\tilde{a}_n(\theta)$, $\tilde{b}_n(\theta)$ and $c_n(\theta)$. Then the Schuster method is to choose that value of θ as an estimator $\tilde{\omega}$ of ω for which $|c_t(\theta)|^2$ is a maximum and the corresponding values of $\tilde{a}_n(\theta)$ and $\tilde{b}_n(\theta)$ as the estimators of a and b in (i).

Two sequences $\{Y_n\}$ and $\{Z_n\}$ of random variables are said to be stochastically equivalent if for any $\varepsilon > 0$

$$\lim \Pr(|Y_n - Z_n| > \varepsilon) = 0.$$

Let ε_t be normal and suppose the assumptions (i) and (ii) hold. Then the periodogram estimators $\tilde{a}_n(\theta)$, $\tilde{b}_n(\theta)$ and $(\theta) = \tilde{\omega}_n$ of the Schuster method are stochastically equivalent to the maximum-likelihood estimators $\hat{a}_n(\theta)$ $\hat{b}_n(\theta)$ and $\hat{\omega}_n$.

The above proposition holds if maximum-likelihood is replaced by least-squares when normality is dropped and the independent random variables ε_t are assumed to have two finite moments with continuous distributions.

A matrix is given which provides an approximation, or more precisely the lower bound, for the matrix of the moments of the maximum likelihood method; it is also an approximation for n large for the matrix of the moments of $\tilde{a}$, $\tilde{b}$ and $\tilde{\omega}$ given by the Schuster method.

The lower Cramér-Rao bound for the variance of the estimator of the period is given and lastly the following is proved: if (i) to (iii) hold and the ε_t are normal with mean zero and finite variance and the parameters (ω, a, b, σ^2) belong to a bounded non-degenerate set of the 4-dimensional Euclidean space, then the maximum likelihood estimators, and hence also the Schuster estimators, are consistent.

2/761 (R. Pro)

THEODORESCU, R. (Institute of Mathematics, Bucharest) **10.1** (1.0)
On chains with complete connections—*In Roumanian*
Studii cercetări matematică (Cluj) (1960) **11**, 195-215 (49 references)

The notion of a chain with complete connections called " chains of infinite order " by other authors, has been introduced into probability theory by Onicescu & Mihoc in their work " Sur les chaînes de variables statistiques " [*C. R. Acad. Sci. Paris* (1935) **200**, 511-512]. This notion, which generalises that of a Markoff chain, has been later extended by Doeblin and Fortet; the systematic study of these chains is due especially to the Roumanian school of probabilists.

The material concerning the theory of chains with complete connections may be found scattered among an entire series of memoirs and monographs. The aim of the present expository paper is to put into perspective the central results of this theory, emphasising the Roumanian contributions [for details, see the recent monograph " Processes with complete connections " by Ciucu & Theodorescu. See also abstract No. **2**/746].

(M. Iosifescu)

2/762

THEODORESCU, R. (Institute of Mathematics, Budapest) **10.1** (1.0)

An equation occurring in the study of stochastic processes—*In Russian*

Rev. Math. Pures Appl. (1960) **5**, 719-722 (4 references)

In this short paper a special functional equation occurring in the theory of processes with complete connections is studied.

Under adequate conditions the solution exists, is unique and represents a probability. For details on these processes, see the recent monograph " Processes with complete connections " by Ciucu & Theodorescu. See also abstract No. **2**/746.

(M. Iosifescu)

2/763

THOMASIAN, A. J. (University of California, Berkeley) **10.5** (-.-)

An elementary proof of the AEP of information theory—*In English*

Ann. Math. Statist. (1960) **31**, 452-456 (5 references)

The author develops some limit relationships among certain characteristics of the sequence of random variables—$(1/n) \log p$ for an arbitrary, not necessarily ergodic or stationary, information source. These permit an elementary combinatorial proof of the asymptotic equipartition property that the sequence from a stationary ergodic source converges in probability to a constant. This was introduced and first proved by McMillan, " The basic theorems of information theory " [*Ann. Math. Statist.* (1953) **24**, 196-219].

(W. J. Hall)

2/764

WALKER, A. M. (Statistical Laboratory, University of Cambridge) **10.7** (10.2)
Some consequences of superimposed error in time series analysis—*In English*
Biometrika (1960) **47**, 33-43 (7 references, 2 tables)

In this paper a study is made of some of the problems that occur when there is a superimposed error in a time series. It is assumed that the errors are additive and form an independent stationary series which is independent of the original time series. The observations are equally spaced and consist of the sum of a stationary linear autoregressive process of order p and a completely random series.

Assuming that each of these two processes, and therefore the sum, are normal then the behaviour of the observed series is completely determined by its autocovariance function. It is easily shown that the effect of the superimposed error is to reduce all the autocorrelations in the same ratio.

Two methods of estimating the parameters of the sequence are considered. The first (*a*) uses the first $p+2$ serial covariance while the second (*b*) uses a different set of $p+2$ consecutive serial covariances. The questions posed are:

(i) What is the best set of covariances to be used in the second method judging by the (asymptotic) efficiency of the estimates?
(ii) How does the efficiency of the second method compare with the first?
(iii) How far are the previous answers affected by non-normality of the processes?

The calculations are very tedious and a detailed examination is given only for the case $p = 1$. Asymptotic efficiencies of the estimates are found and tabulated. Finally the effect for non-normality is considered when it is shown that the non-normality of the error term produces an effect which can, in principle, be corrected but leads to difficult algebra even for the case $p = 1$.

(C. Burrows)

2/765

WORKING, H. (Stanford University, California) **10.7** (6.4)
Note on the correlation of first differences of averages in a random chain—*In English*
Econometrica (1960) **28**, 916-918

In this note the author, after pointing out the importance of bearing in mind the fact that with serial correlation the use of averages can introduce correlations which were not present in the original series, considers the effect of averaging successive groups of items contained in a random chain. These are the most primitive type of stochastic series which have a close resemblance to either stock prices or to certain commodity prices. The author uses an example from his paper " A random-difference series for use in the analysis of time series " [*J. Amer. Statist. Ass.* (1934) **29**] to show that with the number of items in each successive segment only moderately large, the variance of the first differences between averages over successive segments of a random chain is approximately two-thirds of the variance of the first differences between correspondingly positioned terms in the chain.

When the number of items is fairly small, the expected first-order serial correlation of the first differences between average terms in the random chain approximates to $\mathscr{E}(r_1) = +\frac{1}{4}$. Serial correlation of order higher than the first remain zero for first differences of averages of successive groups of terms in the random chain.

(W. R. Buckland)

2/766

ZUBRZYCKI, S. (Math. Inst., Polish Academy of Sciences, Wrocław) **10.2** (8.1)

Concerning plane sampling—*In English*

Second Hungarian Mathematical Congress (1961) (6 references)

Let $y(p)$ be a plane stationary continuous stochastic process depending on a point p in a plane. How to choose n points in a given domain D in order to minimise the mean-square error of estimating the mean value $\eta(D)$ of the process $y(p)$ over D by the average of the values of the process $y(p)$ observed in the chosen points is the question posed by the author. It is the correlation function of the process $y(p)$ which determines the answer to that question. In practice, however, the correlation function is usually not known, and even if we knew it, the computations would be prohibitive. This suggests searching for optimal methods of sampling in some suitably narrowed class of sampling methods with the advantage of having optimal solution for a possibly large class of processes yp.

This paper is a report on some recent investigations in this field: see also Zubrzycki [" Remarks on random stratified and systematic sampling in a plane," *Colloq. Math.* (1958) **6**, 251-264]; Dalenius-Hájek-Zubrzycki [" On plane sampling and some related geometrical problems," in the press] and Ludmilla Zubrzycka [" On the distribution of sampling points in a plane ", *Zastosowania Mat.* (1960) **5**, 161-171: abstracted in this journal No. **2**/719, **8.0**].

(S. Zubrzycki)

2/767

2/768

BÁNKÖVI, G. (Math. Inst., Hungarian Academy of Sciences, Budapest) **11.7** (1.4)

Evaluation of integrals by Monte Carlo methods based on the one-dimensional random space filling—*In English*

Publ. Math. Inst. Hung. Acad. Sci. (1960) **5**, 339-352 (8 references)

The paper deals with the approximative evaluation of integrals of special type

$$\mathscr{I}\{g\} = \int_0^\infty \exp\left[\int_0^z g(u)\, du\right] dz$$

by sampling methods. This procedure is based on the one dimensional random space-filling problem solved by Rényi [*Publ. Math. Inst. Hung. Acad. Sci.* (1958) **3**, 109-127: abstracted in this journal No. **1**/18, **1.4**]. This consists of placing at random disjoint intervals on a longer interval, and determining the mean value of the number of intervals which can be placed. It leads to a similar type of integral.

The author generalised this problem by using intervals whose length is variable, weighting these and changing the placing procedures.

(Ilona Palásti)

2/769

BAZLEY, N. W. & DAVIS, P. J. (National Bureau of Standards, Washington, D.C.) **11.7** (10.1)

Accuracy of Monte Carlo methods in computing finite Markoff chains—*In English*

J. Res., Nat. Bur. Stand. B (1960) **64**, 211-215 (5 references, 3 tables, 2 figures)

Since Monte Carlo methods are generally applied to problems that are too difficult for other types of analysis there is seldom opportunity to compare their results with " exact solutions ". This paper reports results of an experiment performed on the children's game " Chutes and Ladders ", a game which can be interpreted as a finite Markoff chain with 82 states and with given transition probabilities. Statistics, such as the average length of play, are computed on the IBM 704 from 2^{14} simulated plays of the game. These Monte Carlo results are then compared with the " exact solution " obtained by powering the matrix of transition probabilities.

The random element in the game of Chutes and Ladders arises from the toss of a dice for each move. This was simulated by employing the pseudo-random number generator of Taussky & Todd. The unit interval was divided into six equal subintervals and the result of the jth dice throw was identified with the sub-interval in which the pseudo-random number t_j fell. The average length of play after 16,384 games was 39,420 throws. The exact answer is 39,224.

The author concluded that the results of Monte Carlo calculations are in agreement with the theoretical results obtained from the matrix calculations and the convergence according to the $1/\sqrt{N}$ law.

Note: In equation (5.2) " σ " should read " σ^2 ".

(H. H. Ku)

2/770

DEMING, Lola S. (Statist. Eng. Lab., National Bureau of Standards, Washington) **11.9** (6.0)
Selected bibliography of statistical literature, 1930 to 1957.
I. Correlation and regression theory—*In English*
J. Res., Nat. Bur. Stand. B (1960) **64**, 55-68

This is the first in a series of bibliographies dealing with various specific subjects in the field of statistics. The series is intended to be of service by listing the most important contributions in statistical theory and method, as judged by two prominent reviewing journals [*Zentrablatt für Mathematik* for the years (1930) to (1939), and *Mathematical Reviews* from (1940) onward]. The scheme of classification follows that of the *Mathematical Reviews* Annual Index.

The titles are given exactly as in the reviewing journal, and are listed under the first-named author only in the case of joint authorship. Joint authorship is denoted by a symbol preceding the surname and the authors are cross-referenced to the first-named author.

In addition to the name of publication, volume, initial page number, and date of publication, the volume and page numbers are given for the two reviewing journals in which the abstract appears.

There are approximately 520 entries in this bibliography on correlation and regression theory. See also II. " Time Series "; abstracted in this journal No. **2**/772, **11.9**: III. " Limit Theorems "; abstracted in this journal No. **2**/773, **11.9** and IV. " Markoff chains and stochastic processes "; abstracted in this journal No. **2**/774, **11.9**.

(H. H. Ku)

2/771

DEMING, Lola S. (Statist. Eng. Lab., National Bureau of Standards, Washington) **11.9** (10.0)
Selected Bibliography of statistical literature, 1930 to 1957.
II. Time Series—*In English*
J. Res., Nat. Bur. Stand. B (1960) **64**, 69-76

This is the second in a series of bibliographies dealing with various specific subjects in the field of statistics. References and titles of important contributions to the study of time series have been taken from a wide variety of technical journals published in the many languages and countries which have been actively engaged in statistical analysis. For a description of the bibliography in general see I. " Correlation and Regression Theory ", abstracted in this journal No. **2**/771, **11.9**.

There are approximately 280 entries in this bibliography on time series. See also III. " Limit Theorems "; abstracted in this journal No. **2**/773, **11.9** and IV. " Markoff Chains and Stochastic Processes "; abstracted in this journal No. **2**/774, **11.9**.

(H. H. Ku)

2/772

DEMING, Lola S. (Statist. Eng. Lab., National Bureau of Standards, Washington) **11.9** (1.5)
Selected Bibliography of statistical literature, 1930 to 1957.
III. Limit theorems—*In English*
J. Res., Nat. Bur. Stand. B (1960) **64**, 175-192

This is the third in a series of bibliographies dealing with various specific subjects in the field of statistics. References and titles of important contributions concerning limiting distributions have been taken from technical journals published throughout the world since 1930.

There are approximately 670 entries in this bibliography. The two preceding ones are: I. " Correlation and Regression Theory "; abstracted in this journal No. **2**/771, **11.9**: II. " Time Series "; abstracted in this journal No. **2**/772, **11.9** and there is also IV. " Markoff Chains and Stochastic Processes " ; abstracted in this journal No. **2**/774, **11.9**.

(H. H. Ku)

2/773

DEMING, Lola S. & GUPTA, D. (Statist. Eng. Lab., Nat. Bur. Stand. & Catholic Univ., Washington) **11.9** (10.0)
Selected Bibliography of statistical literature, 1930 to 1957.
IV. Markoff chains and stochastic processes—*In English*
J. Res., Nat. Bur. Stand. B (1961) **65**, 61-93

This is the fourth in a series of bibliographies dealing with various specific subjects in the field of statistics. Titles and references of important contributions to the study of Markoff chains and stochastic processes have been taken from technical journals published since 1930 in the many countries that have been actively engaged in statistical analysis.

There are approximately 1250 entries in this bibliography. See also I. " Correlation and Regression Theory "; abstracted in this journal No. **2**/771, **11.9**: II. " Time Series"; abstracted in this journal No. **2**/772, **11.9** and III. " Limit Theorems "; abstracted in this journal No. **2**/773, **11.9**.

(H. H. Ku)

2/774

EVANS, Marilyn J. & HILL, G. W. (Div. Math. Statist., CSIRO, Adelaide & Melbourne) **11.1** (3.1)
Tables of $z = \text{arc tanh } (r)$—*In English*
Div. Math. Statist. Tech. Paper No. 7 (1961) 13 pages (2 references, 5 tables)

This technical paper supplements Table VII of Fisher & Yates [*Statistical Tables for Biological, Agricultural and Medical Research*, 4th Edn. (1953) Edinburgh: Oliver & Boyd]. Tables are given for the function

$$z = \text{arc tanh } (r) = \tfrac{1}{2}\{\ln(1+r) - \ln(1-r)\}$$

required to obtain Fisher's z-statistic from known values of r.

There is a brief introduction to the tables, followed by:

Table 1 (*a*): values of z for $r = 0\ (0{\cdot}001)\ 1$,
Table 1 (*b*): values of z for $r = 0{\cdot}9\ (0{\cdot}0001)\ 1$,
Table 2 (*a*): values of z for $r^2 = 0\ (0{\cdot}0001)\ {\cdot}05$,
Table 2 (*b*): values of z for $r^2 = 0\ (0{\cdot}001)\ 1$,
Table 2 (*c*): values of z for $r^2 = 0{\cdot}9\ (0{\cdot}0001)\ 1$.

(J. Gani)

HAIGHT, F. A. (University of California, Berkeley) **11.9** (2.0)
Index to the distributions of mathematical statistics—*In English*
J. Res., Nat. Bur. Stand. B (1961) **65**, 23-60

A fairly complete index of references to results on statistical distributions published before January 1958 is presented. The distributions covered are characterised as normal, type III, binomial, discrete, distributions over (a, b), distributions over (a, ∞), distributions over $(-\infty, \infty)$, miscellaneous univariate, miscellaneous bivariate, and miscellaneous multivariate. The number of entries varies from one or two for less well-known distributions to several hundred for the normal distributions. This index should serve to eliminate unnecessary derivation of results already in the literature.

The material given under each distribution consists of a number of entries, organised under a standard order of headings for well-known distributions with sufficient number of entries, for example, the normal:

A. Functions and Parameters,
B. Derived distributions.
C. Estimation.
D. Testing statistical hypotheses.
E. Miscellaneous.

Direct and indirect references are used in this index: direct references include twenty-six coded journals and eighteen coded books, and uncoded literature. The indirect references give reviews by *Mathematical Reviews* and *Zentralblatt für Mathematik*—usually for papers that appear in obscure sources. A distribution is included if its density, or probability function is a known, explicit function, and an entry must exhibit a property of the distribution in question. Exceptions include certain densities specified only implicitly or in terms of their cumulative probability function, or characteristic function. Historical information, important applications, and bibliographies are sometimes given. An alphabetical index to the index to distributions is also included.

An example of the entries is:

Wilk's distribution of dispersion determinant, etc.: (*e*) 3:26 where the reference given is *Sankhyā*, Vol. 3, p. 26.

(H. H. Ku)

NAGAR, A. D. (Netherlands School of Economics, Rotterdam) **11.7** (6.6)

A Monte Carlo study of alternative simultaneous equation estimators—*In English*

Econometrica (1960) **28**, 573-590 (4 references, 8 tables)

In this paper the author discusses a Monte Carlo approach to four methods of estimating the small sample properties of parameters of simultaneous equations. The methods discussed are:

- (i) least squares
- (ii) two-stage least squares
- (iii) unbiased second-moment
- (iv) minimum second-moment.

After a reference to a paper by Wagner " A Monte Carlo study of estimates of simultaneous linear structural equations " [*Econometrica* (1958) **26**, 117-133], the author states that his paper is a further examination of Wagner's models. This examination is conducted in the light of the alternative procedures now available in order to carry out a comparison. The present author is able to disregard the methods of estimation already discussed by Wagner. After stating that the two equations used by the previous author are " extreme ", in a certain sense, the present author hopes that a further examination may give valuable information regarding these " extreme " type equations. This paper considers both the over-identified and just-identified equation in the model.

The second section deals with the alternative estimation procedures; Wagner himself having considered the limited-information estimates of the over-identified equations. The procedures and their properties are described but not proved.

After a section which gives further consideration to Wagner's models the author gives numerical results for two models in terms of one hundred samples of size twenty. The four methods of estimation are considered, all belonging to the class, and the final section of this paper gives results which include observations on sampling variance, bias and asymptotic standard error. The author gives eight tables of various points of comparison between the methods.

(W. R. Buckland)

2/777

THEIL, H. (Netherlands School of Economics and Stanford Univ.) **11.8** (6.4)

Best linear index numbers of prices and quantities—*In English*

Econometrica (1960) **25**, 464-480 (5 references, 1 figure)

This paper deals with the statistical type of approach to the subject of index numbers. It differs from the usual approach in as much as it is not necessarily only concerned with one single pair of periods. The author considers index numbers for an arbitrary number of periods which are determined by the prices and quantities of all separate periods. This procedure is of course necessarily *ex post* as it can only be used when the prices and quantities for the whole year are available; it is also suitable for computation of prices and quantity indices needed for an econometric model. Another situation to which it is applicable is the comparison of prices and quantities between different geographical units and the symbol t in the matrix used in the section on best linear price index numbers can also be used to denote either time units or geographical units, although the exposition is largely confined to variations over time.

The primary sections of this paper deal with the geometric and algebraic aspects of this approach which is a " linear " one in the sense that the price index vector is a linear function of the matrix of the individual prices. It is closely related to principal component analysis from which it differs by virtue of the fact that it considers two sets of variables, both prices and quantities. Sections two and three deal with prices only and are followed by a completely analogous approach to quantities in section four.

In the fifth section price and quantities are combined and the following section deals with a filtering index which serves to measure the statistical quality of the price and quantity index vectors. Partial index numbers and problems of aggregation follow in the next section and in conclusion the author comments on low filtering indexes and the weighting of procedures to deal with years which are considered " outlying ".

(W. R. Buckland)

2/778

THEIL, H. (Netherlands School of Economics, Rotterdam) **11.8** (3.4)
Best linear index numbers of prices and quantities—*In English*
Econometrica (1960) **28**, 464-480 (5 references, 1 figure)

This paper is statistical in that it is concerned with the optimum specification of a central tendency for price and quantity ratios. Its approach is different, however, in that it is not necessarily concerned with one pair of periods but with an arbitrary number of complete periods. This approach is suitable for the estimation of an econometric model and can be applied on a current basis for price and quantity comparisons of different geographical units.

In the second section the approach of the paper is demonstrated geometrically followed by an algebraic account in section three. The approach is linear in the sense that the price index is a linear function of the matrix of individual prices. Because a quadratic form relating to certain discrepancies is minimised the method of this paper is a " best " approach and closely related to principal component analysis. While sections two and three deal only with prices, the analogous situation for quantities is dealt with in section four. The two situations are combined in sections five and six. The next section deals with a " fitting-index " which measures the statistical quality of the price and quantity vectors. The problems of partial index numbers and problems of aggregation are dealt with in section eight.

The final section contributes some general comments by way of interpretation. For example, if the " fitting-index " is low then the second largest latent root of the price and quantity vectors is used: both being pairwise. If any particular year is an " outlying " observation it is possible to apply a weighting procedure resembling Aitken's method of generalised least squares [*Proc. Roy. Soc. Edinb.* (1934-35) **55**, 42-48].

(W. R. Buckland)

2/779

VENEKAMP, P. E. (University of Amsterdam) **11.0** (-.-)
Descriptive statistics—*In Dutch*
Statist. Neerlandica (1961) **15**, 171-174

In this article the concepts of descriptive statistics and theoretical statistics are compared. It is concluded that descriptive statistics are to be considered an applied science, aimed at presentation and interpretation of figures. In this context the methods of theoretical statistics are also to be applied.

(P. E. Venekamp)

2/780